

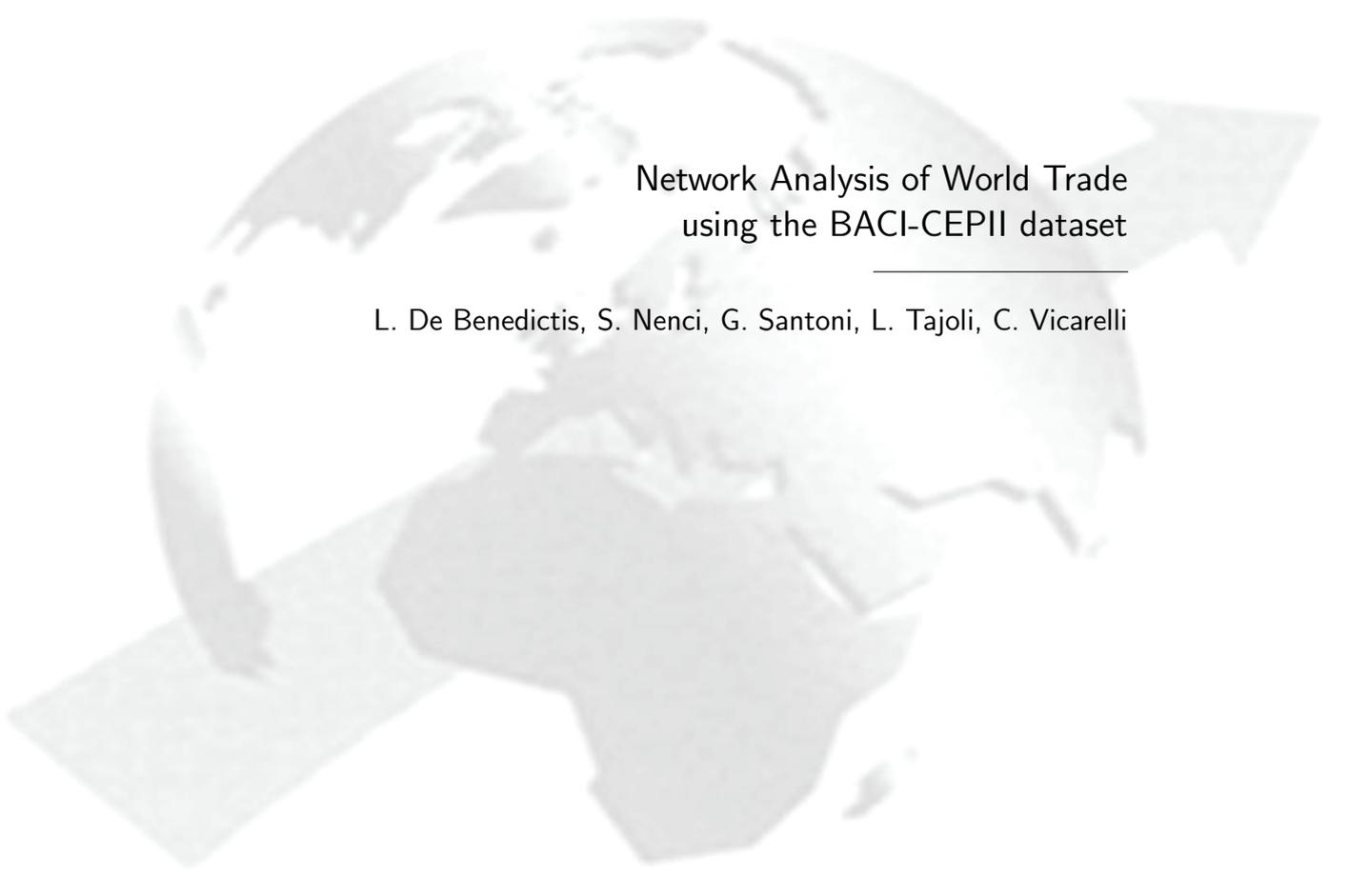


**CEPII**

CENTRE  
D'ÉTUDES PROSPECTIVES  
ET D'INFORMATIONS  
INTERNATIONALES

No 2013 – 24  
August

DOCUMENT DE TRAVAIL



Network Analysis of World Trade  
using the BACI-CEPII dataset

---

L. De Benedictis, S. Nenci, G. Santoni, L. Tajoli, C. Vicarelli

---

**Table of contents**

Highlights and Abstract . . . . .	3
Points clefs et résumé . . . . .	4
1. Introduction . . . . .	5
2. Visualizing World Trade . . . . .	9
2.1. Graphs and Networks: Basics . . . . .	15
2.2. Graphs and networks: sociograms . . . . .	17
2.3. Zooming in and out trade data: Local and Global views . . . . .	21
3. Local and global centrality measures and relative countries' position . . . . .	22
3.1. Centrality measures . . . . .	23
3.2. Global trade through centrality indexes lenses . . . . .	33
4. Sectoral Trade Networks: Bananas, Cement, Movies, Oil, Footwear and Engines . . . . .	39
4.1. Bananas . . . . .	39
4.2. Cement . . . . .	41
4.3. Movies . . . . .	43
4.4. Crude Oil . . . . .	43
4.5. Footwear . . . . .	45
4.6. Engines . . . . .	46
4.7. Centrality in sectoral data . . . . .	48
5. Conclusions . . . . .	50
6. Appendix: Data and indicators . . . . .	52
7. Appendix: R, Stata and Pajek routines . . . . .	53
7.1. Appendix: Calculating networks cumulative distributions in R . . . . .	53
7.2. Appendix: A brief note on Stata routines for network centrality measures . . . . .	54
7.3. Appendix: Pajek routines for network visualization . . . . .	54



## Network Analysis of World Trade using the BACI-CEPII dataset

L. De Benedictis, S. Nenci, G. Santoni, L. Tajoli, C. Vicarelli

### Highlights

- This paper explores the World Trade using the Network Analysis and introduces the reader to some of the techniques used to visualize, calculate and synthetically represent network trade data.
- The paper shows different visualizations of the network and describe its topological properties, producing and discussing some of the commonly used Network's statistics, and presenting some specific topics.
- All in all, this paper shows that Network Analysis is a useful tool to describe bilateral trade relations among countries when interdependence matters, and when trade relations are characterized by high dimensionality and strong heterogeneity.

### Abstract

In this paper we explore the BACI-CEPII database using Network Analysis. Starting from the visualization of the World Trade Network, we then define and describe the topology of the network, both in its binary version and in its weighted version, calculating and discussing some of the commonly used network's statistics. We finally discuss some specific topics that can be studied using Network Analysis and International Trade data, both at the aggregated and sectoral level. The analysis is done using multiple software (Stata, R, and Pajek). The scripts to replicate part of the analysis are included in the appendix, and can be used as an hands-on tutorial. Moreover, the World Trade Network local and global centrality measures, for the unweighted and the weighted version of the Network, calculated using the bilateral aggregate trade data for each country (178 in total) and each year (from 1995 to 2010,) can be downloaded from the CEPII webpage.

*JEL Classification:* F10

*Keywords:* International trade, Network Analysis, Density, Centrality, Stata, R, Pajek.



## Network Analysis of World Trade using the BACI-CEPII dataset

L. De Benedictis, S. Nenci, G. Santoni, L. Tajoli, C. Vicarelli

### Points clefs

- Ce document applique la méthode de l'analyse des réseaux aux flux du commerce international et introduit le lecteur aux techniques que cette méthode utilise pour visualiser, calculer et représenter de façon systématique les réseaux du commerce international.
- Il présente différentes visualisations des réseaux du commerce international et décrit leurs propriétés topologiques.
- Ce document montre que l'analyse des réseaux est très utile à la description et à la compréhension des interdépendances provenant des relations commerciales bilatérales ainsi qu'à l'analyse de flux commerciaux à plusieurs dimensions et forte hétérogénéité

### Résumé

Dans ce document de travail, nous exploitons la base de données du commerce international BACI, construite au CEPII, en utilisant la méthode de l'analyse des réseaux. Partant de la visualisation du Réseau Mondial du Commerce International, nous définissons et décrivons la topologie de ce réseau, à la fois dans sa version binaire et dans sa version pondérée, puis nous examinons les indicateurs statistiques du réseau couramment utilisés. Nous discutons enfin des sujets spécifiques qui peuvent être étudiés en appliquant l'analyse des réseaux au commerce international, aux niveaux agrégé et sectoriel. L'analyse est effectuée à l'aide de différents logiciels (Stata, R, et Pajek). Les scripts pour reproduire une partie de l'analyse figurent en annexe ; ils peuvent être utilisés pour se familiariser avec ce type d'analyse. Par ailleurs, les mesures de centralité mondiale et locale du Réseau Mondial du Commerce, dans les versions non pondérée et pondérée du réseau, calculées en utilisant l'ensemble des données bilatérales de chaque pays et pour chaque année (de 1995 à 2010), peuvent être téléchargées à partir du site web du CEPII.

*Classification JEL* : F11, F14

*Mots clés* : Commerce International, Analyse des Réseaux, Densité, Centralité, Stata, R, Pajek

## Network Analysis of World Trade using the BACI-CEPII dataset<sup>1</sup>

Luca De Benedictis\*

Silvia Nenci<sup>†</sup>

Gianluca Santoni<sup>‡</sup>

Lucia Tajoli<sup>§</sup>

Claudio Vicarelli<sup>¶</sup>

### 1. Introduction

At the beginning of this century, celebrating the centennial of Bertil Ohlin's birth, Donald Davis and David Weinstein, tracing the evolution of the research in international trade, were praising for an empirical shift of the discipline (Davis and Weinstein, 2001).<sup>2</sup> In ten years this shift is made.

The fundamental ingredients of this new empirical wave in international trade analysis are the cross-fertilization of different field in economics, especially on empirical grounds; the diffusion of econometrics in the background studies of the new generation of international trade economists; and even more fundamental, the increasing availability of data at the macro, sectoral and firm level suited for the analysis of international trade issues. International institutions play an essential role in this respect. The United Nation system (UN), the World Bank, the International Monetary Fund, the Eurostat system, among others, collect, produce and diffuse an enormous amount of internationally comparable

---

<sup>1</sup>This work has received financial support by the Italian Ministry of Education, University and Research under the grant PRIN 2009 "The international trade network: empirical analyses and theoretical models" ([www.netrade.org](http://www.netrade.org)). We would like to thank Giorgio Fagiolo, Massimo Riccaboni, Stefano Schiavo, Gianluca Orefice for the useful comments and Matthieu Crozet for the encouragement. All errors are our responsibility.

\*DED, University of Macerata, Via Crescimbeni 20, Macerata (MC), Italy, ([debene@unimc.it](mailto:debene@unimc.it))

<sup>†</sup>DE, University of Roma Tre, Rome (RM), Italy, ([silvia.nenci@uniroma3.it](mailto:silvia.nenci@uniroma3.it))

<sup>‡</sup>IE, S.Anna School of Advanced Studies, Pisa (PI), Italy, ([g.santoni@sssup.it](mailto:g.santoni@sssup.it))

<sup>§</sup>Politecnico di Milano, Milan (MI), Italy, ([lucia.tajoli@polimi.it](mailto:lucia.tajoli@polimi.it))

<sup>¶</sup>ISTAT, Rome (RM), Italy, ([cvicarelli@istat.it](mailto:cvicarelli@istat.it))

<sup>2</sup> Davis and Weinstein's final words are worth quoting: "*Crisp, lucid theory will always play a central role in the field. But this needs to be complemented by a serious encounter with data.*" "*It is time for each international economist to accept the challenge to make empirical analysis a central feature of our work and dialogue. We have a world to discover.*"

data. This is a gold mine for empirical analysis. Restricting the focus to the specific data we are going to work with, bilateral trade flows are constantly collected at a very detailed commodity level (6 digit Harmonized System) by the United Nations Statistical Department, in the *ComTrade* database.<sup>3</sup>

Some research centers have recently started to produce their own variants of the original database. The French research center CEPII (*Centre d'Etudes Prospectives d'Informations Internationales*) is one of the first that started offering to the public the outcome (or spin-off) of its research in terms of data. The BACI (*Base pour l'Analyse du Commerce International*) dataset is one of these outcomes. The original motivation for producing BACI is that, in spite of the richness of the data reported in ComTrade, the attempt to account for highest number of countries, the largest period of time and the most disaggregated product level is plagued by the too many missing flows in the original UN database. BACI uses a *mirror statistic* strategy to impute missing data (see Gaulier and Zignago (2010) for a description of the data, the methodology used, and the comparison with other databases). In general, the same 6-digit HS trade flow should be reported by both the exporting country  $i$ , as exports, and by the importing country  $j$ , as imports. Both figures can be present, or missing, or just one of the two can be missing. In this last case, BACI imputes the missing data using the mirror statistic reported by the trade partner. Using a 'reconciliation' methodology,<sup>4</sup> (Gaulier and Zignago, 2010) BACI reduces the number of missing values substantially, the remaining ones being those between two non-reporting countries.

The BACI database is used overwhelmingly in applied trade analysis.<sup>5</sup> It is rapidly diffusing among scholars, it is constantly updated by the CEPII and it is largely available. In this

---

<sup>3</sup> The UN *ComTrade* database is the largest depository of international trade data available online. It provides international (imports and exports) merchandise trade statistics for about 200 reporting countries from 1962 to 2012. The database contains detailed statistics reported by national statistical agencies, and standardized by the UN Statistics Division, using the UN/OECD CoprA internal processing system. Commodities are reported in the current classification and revision (HS2002 in most cases) and are converted all the way down to the earliest classification SITC revision 1. Commodities are therefore classified according to SITC (Rev.1 from 1962, Rev.2 from 1976 and Rev.3 from 1988), the Harmonized System (HS) (from 1988 with revisions in 1996, 2002 and 2007) and Broad Economic Categories (BEC). Data do not include flows below 1,000 US dollars. Further information is included in the webpage of the UN ComTrade statistics: <http://unstats.un.org/unsd/tradekb/Knowledgebase/Comtrade-User-Guide>.

The IMF data would not be analyzed in this paper, however it is a feasible alternative. Just for comparison, the IMF *Direction of Trade Statistics* database contains annual, quarterly and monthly time series on merchandise imports and exports for 187 countries from 1981 to May 2012 (the most recent monthly data available). The *Historical Direction of Trade Statistics* database covers years from 1948 to 1980.

<sup>4</sup> The methodology forces the reported flow to be coherent when observed as import and as export. The resulting data is therefore not suited for the analysis of trade smuggling based on the information derived from the discrepancy between reported flows (Fisman and Wei, 2004, 2009; Javorcik and Narciso, 2008; Rotunno and Vézina, 2012).

<sup>5</sup> In December 2012 there were 451 papers quoting the BACI-CEPII database in [scholar.google.com](http://scholar.google.com).

paper we explore the BACI dataset through the lenses of Network Analysis, visualizing the World Trade Network, defining and describing the topology of the network, producing and discussing some of the commonly used Network's statistics, and presenting some specific topic that can be studied using the Network Analysis of relational or dyadic ( $ij$ ) aggregated and sectoral international trade data. The analysis is done using multiple software (Stata, R, and Pajek), and the scripts to replicate some sections of the analysis are included in the appendix, and can be used as an hands-on tutorial to the network analysis of trade flows. Moreover, the World Trade Network local and global centrality measures, for the unweighted and the weighted version of the Network, calculated using the bilateral aggregate trade data for each country (178 in total) and each year (from 1995 to 2010) can be freely downloaded from the CEPII webpage.

Before entering in the details of the analysis, it is appropriate to answer a preliminary question: why do we want to look at international trade data (as in the BACI-CEPII database) using Network Analysis? What is so special about Network Analysis with respect to other descriptive statistics generally used in applied international trade analysis (Bowen, Hollander and Viaene, 2012)? The short answer is that, first of all, networks are about *relations*. The fundamental piece of information of networks is the dyad  $ij$ ; not the monads  $i$  or  $j$ , but the relationship between them. However, the specificity of networks is that the relation between  $i$  and  $j$  is not analyzed in isolation, but it is studied focusing on its *structural dimension*, that is taking into account the effect of  $z$  in the relation between  $i$  and  $j$ . Extending the *effect of others*, or in our case the “ $z$  country effect,” to the many  $z$  included in the set of possible trade relations, the resulting image is a network in its essence. The implication of this “structural view” is that the relation between  $i$  and  $j$  cannot be considered independent from the relation between  $i$  and  $z$ , and between  $j$  and  $z$ . Therefore the characteristic of *interdependence* is the hinge of networks.

In general, the *effect of others* could be reduced eventually to the “average effect of others,” as in the studies on *peer effects*. In that case, the implicit assumption is that the peers  $z$ s are somehow homogeneous, so that the mean is a meaningful central statistic of the distribution of peers characteristics. However, this is generally not the case in social networks, which are instead characterized by a high degree of *heterogeneity* and a power law distribution of the topological properties of the network's nodes. We will quantify the country heterogeneity in the World Trade Network through the use of *centrality* measures. This heterogeneity makes the case for the full structural analysis of the complete set of relations in the network. This is what we will do in the next pages.

In one line, if the *effect of others* on self is of marginal interest, there is no point of bothering with network analysis. If on the contrary, structural interdependence and third part's ( $z$ 's) effects are of interest, here network analysis can be a fruitful complement to more traditional research.

Network Analysis has rapidly evolved in these recent years.<sup>6</sup> From the analysis of small communities undertaken by sociologists and anthropologists, to the very large networks studied by physicists and computer scientists; to the focus on Graph Theory mathematical models, to the empirical analysis of networks. The increasing availability of software tools for the analysis, simulation and visualization of network data is boosting the empirical side of the discipline.<sup>7</sup> Only mentioning the software used in subsequent analysis, the Stata Graph Library (SGL) has just been released by Miura (2012). It is suited for *directed* or *undirected* one-mode networks containing *arcs* that are either *unweighted* (i.e. bilateral trade partnership, in our case) or *weighted* (i.e. bilateral trade flows). In spite of the network jargon that will be clarified in a little while, those characteristics are the ones of the BACI trade data we will work with. Therefore we will make use of SGL in calculating the network statistics discussed in the paper. Computational inefficiency makes SGL impractical for large networks. In this paper, the graph associated to the World Trade Network  $\mathcal{G} = (\mathcal{V}, \mathcal{L})$  has an average dimension<sup>8</sup> of 178 vertices ( $\mathcal{V} = 178$ ) and 22,000 trade links ( $\mathcal{L} = 22,000$ ): it can be considered a medium-size network and can be safely operationalized through SGL.

Since a while the R statistical computing environment (R Development Core Team, 2007) witnessed the emergence of different tools for the analysis of relational data. The `igraph` library (Csardi and Nepusz, 2006), the `network` library (Butts, 2008a), the `sna` library (Butts, 2008b), among others, have been collected under the umbrella of the `statnet` library (Goodreau, Handcock, Hunter, Butts and Morris, 2008), while the `tnet` library as been realized expressly do deal with weighted and longitudinal networks (Opsahl, 2009). We use this tools to compare and integrate the results obtained through the SGL Stata routine. The advantages of using R are countless, and the capabilities of the R libraries are far beyond the ones that are highlighted in this paper. The computational efficiency of `igraph` allows to work with very large networks (e.g. 1.8 million relations) and makes it our preferred tool for the production of networks statistics distributions. The `tnet` library as been used to calculate weighted networks statistics and to overcome the limits

---

<sup>6</sup> For an introduction to the evolving discipline of Network Analysis see Scott (2000), for a beautiful recent introductory treatment of the topic see Newman (2010), and for a more technical treatment see Wasserman and Faust (1994). Networks have also entered economic theory in recent years. See Vega-Redondo (2007); Goyal (2009); Jackson (2010) for comprehensive overviews, and Zenou (2012) for a short but illuminating introduction.

<sup>7</sup> The International Network for Social Network Analysis (INSNA) is listing in its webpage around twenty software notified by INSNA Members. It is a large underestimate of the existing options available on the web. For a selective overview of the tools available see Combe, Langeron, Egyed-Zsigmond and Géry (2010).

<sup>8</sup> The BACI-CEPII data used in this paper includes a subset of countries. The more than 200 countries included in the original database is therefore reduced to 178, including only the countries that where reporting trade flows for at least one international partner in *all* years between 1995 and 2010. Some of the measures that will be calculated in section 3 require the square matrix of yearly bilateral trade flows to be full-rank.

of SGL.

Finally, the choice of making use of *sociograms*<sup>9</sup> lead us to take advantage of the graphic capability of *Pajek* (de Nooy, Mrvar and Batagelj, 2011).<sup>10</sup> Data in *Pajek* format can be read by both *SGL* and *igraph* and *statnet*. Therefore the BACI-CEPII dataset has been converted in a *Pajek* format and all the visualizations included in the paper were done with *Pajek*. Instead, the network statistics for the World Trade Network have been calculated with *SGL*, *igraph*, *tnet* and *Pajek*.

This paper offers a synthetic introduction to the application of Network Analysis to international trade data, as collected by the Cepii in the BACI-CEPII database. The main goal of the analysis is to introduce the reader to some of the techniques used to visualize, calculate and synthetically represent network trade data.<sup>11</sup> We especially focus on centrality measures, as a way of describing the level of country heterogeneity in bilateral trade flows. We fully describe the centrality measures of the World Trade Network in 2007. However, the interested reader could find the full set of centrality measures between 1995 and 2010 in the [CEPII databases page](#). Moreover, some of the scripts to calculate some specific centrality measures using *Stata*, *R* or *Pajek*, or to reproduce some of the figures presented in the paper, are available in the paper's Appendix.

The structure of the paper is the following: in section 2 we present different ways to visualize the World Trade Network and present some general characteristics of the network; in section 3 we define various centrality measures, and we derive them for the countries in our sample; in section 4 we move from aggregate data to sectors, describing the structure of the network for some selective products; Finally section 5 concludes.

## 2. Visualizing World Trade

If we look at the contemporary world trade the general impression that we receive is of a significant economic interaction among countries. Bilateral trade flows show intense and wide relations between countries, among which some stronger links emerge. A quite natural visualization of trade flows is through the use of a cartogram and of arrows linking

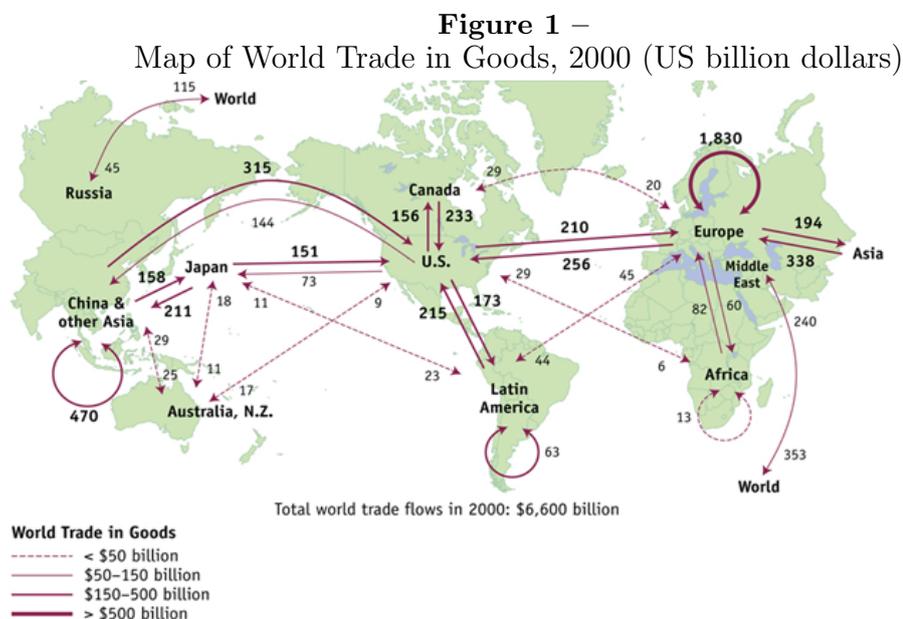
---

<sup>9</sup> What is called a sociogram, is now the classical visual representation of networks. It is a diagram, with individuals (countries in our case) represented by vertices or nodes and their (social) relationship to one another by links or lines or edges or arcs (trade flows in our case). The sociogram was elaborated in the 1930s by the social psychologist Jacob Moreno (Moreno, 1934) and have been given some solid Graph Theory bases by Cartwright and Harary (1977), the founders of modern social network analysis.

<sup>10</sup> *Pajek* (the Slovene word for spider) is a freeware program for the analysis and visualization of very large networks. It is computationally efficient and can be used for the calculus of networks statistics and more.

<sup>11</sup> Being a simple technical introduction to the topic, the paper does not review the growing empirical literature on the application of network analysis to international trade. For an overview of the literature see Garlaschelli and Loffredo (2005), Kali and Reyes (2007), Fagiolo, Reyes and Schiavo (2008) and De Benedictis and Tajoli (2011).

countries or geographical areas. One example is the one offered by Feenstra and Taylor (2008) and reproduced in figure 1. It presents the geographical World map where arrows of



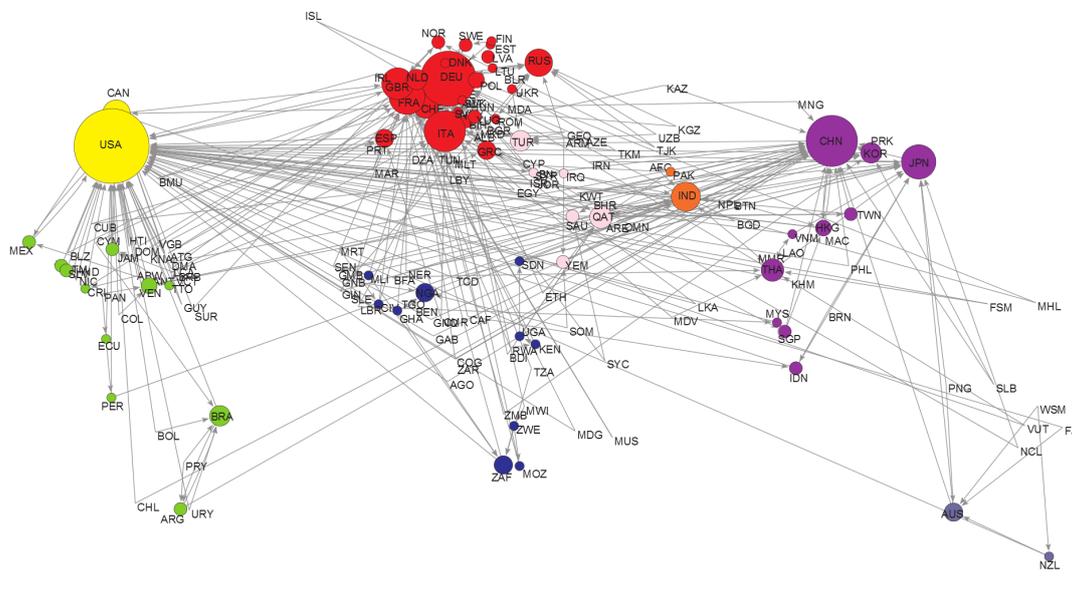
**Source:** Feenstra and Taylor, 2008, Chapter 1.

**Note:** The thickness of links is proportional to the volume of trade. Intra-regional trade flows are depicted by loops.

trade in merchandise goods are superimposed to the cartogram. By definition, the distance between countries in figure 1 matches their geographical distance, and the curvature of the arrows is purely esthetic. The picture is very efficient in showing the intensity of trade links between some countries and continents, highlighting some significant trends: (i) the leading role of Europe in world trade; (ii) the large amount of trade flows at regional level (within Europe, within the Americas and also between Asian countries); (iii) a relevant trade link between the United States and Europe (about 35% of world trade can be accounted for by these two trading partners); (iv) the important share of world trade coming from Asia (exports from it totaled about one-quarter of global trade) and its considerable trade links with both Europe and the United States; (v) the marginal position of Africa (that accounts for only 2.5% of World trade), and its close relationship with European countries.

Although suggestive, this picture does not give full account of the implication of the interdependence among countries. To see why, let's produce a similar figure using a selection of the available information on the full set of bilateral trade flows. Instead of describing a limited number of countries or continents, as in figure 1, let's plot the entire set of 178 countries in our dataset according to geographical coordinates, drawing only the two main export flows for each country. The result is in figure 2.

**Figure 2** –  
Map of World Trade in Goods as a Geographical Network (major two export partners,)  
2007.



**Note:** For each country, only the export flows toward the first and second trade partner are considered. Country labels are the iso3 country codes (see the Appendix for a complete list of the countries included in the analysis). The position of each country-node corresponds to the latitude and longitude coordinates of the centroid of the country main geographical area. The size of the circle associated to each country is proportional to the number of inflows. Different colors correspond to different geographical regions. Trade data come from BACI-CEPII dataset. The network is drawn using Pajek. The sequence of Pajek commands necessary to reproduce the above figure is included in the Appendix.

Each country is represented by a *node* (or a vertex) in the network, labeled with the respective iso3 country code (the entire list of countries, iso3 codes and respective geographical regions is included in the Appendix). Different colors correspond to different continents or sub-continents. Since we are focusing only on the two main export markets for each country, there are only two *links* (or arcs) directed from each country  $i$  toward the two main export partners (e.g. Chile (CHL) is connected to the US (USA) and China (CHN): those are its main export partners, on the other hand, Chile is not among the major two export partners for any of the countries represented in figure 2). The size of vertices is proportional to the number of incoming arcs: a large circle indicate that the corresponding country is among the main partner for a large number of exporting countries.

Apart from the different reference year, and the data-selection strategy adopted to visualize the way too large amount of information included in global bilateral trade, figure 1 and figure 2 are conveying the same image of world trade. The US, Germany (DEU) and China

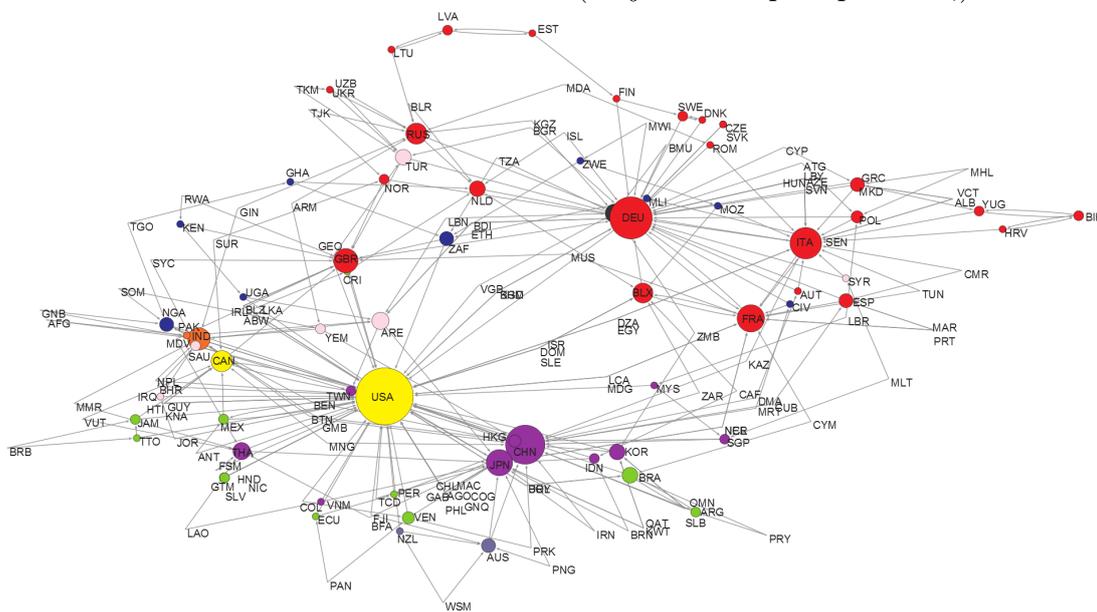
are the reference markets for a large number of countries, regional trade is substantial (e.g. Europe), some continents are largely disconnected (e.g. Africa and South America), other geographical regions seem more connected (e.g. Oceania and East Asia). More generally, the world trade system seems to be characterized by a high level of interconnectivity or interdependence.

As we were saying, the geographical representation of international trade does not allow to describe the effect of this interdependence on each single country. To see why, let's proceed with the following thought experiment. Let's imagine to delete an arc from figure 2: the link between Brazil (BRA) and the US (USA), for example. Nothing special will happen to the whole picture. Just the circle corresponding to the US would become marginally smaller, but from a structural point of view the rest of the picture will remain unaltered: the position of countries, being fixed by geography, is spatially unaffected, and so is the position of the other arcs in the picture. Basically, the other countries' trade is unaffected by the sudden loss of Brazilian exports to the US. To borrow an expression popularized by Anderson and vanWincoop (2003), it is as if the interdependence among countries was only examined at the bilateral level instead of the multilateral level. The same would happen if the link between Brazil and the US would have been substituted by a new link between Brazil and Germany (DEU). Apart from the size of the circles corresponding to the US and Germany nothing else would have changed in figure 2. This is the visual analog of the assumption of conditional independence among dyads imposed on international trade flows.

Instead, if we want countries' interactions to be accounted in determining the *relative position* of each country in the whole trading system, we should get rid of the constraint imposed by a geographical representation of international trade and move from physical space to *topological space*. This is what we do next, using Network Analysis techniques for the visualization of dyadic data.

Figure 3 is obtained relaxing the geographical constraint on the position of countries in the trade network and applying what is called a *force-directed algorithm* (Kamada and Kawai, 1989; Fruchterman and Reingold, 1991; de Nooy *et al.*, 2011) on the same data used to produce figure 2. Colors, labels and the size of the circle of each country-node are the same as before. Being simplistic, the algorithm acts as a balanced spring system that minimizes the energy in the system. In other words, it is as if countries were linked through springs: countries which are connected tend to stay close, while countries which are not connected tend to be placed far apart. However, the position of each country does not depend only on its bilateral links but also on the indirect *effect of others*: the trade partners of its trade partners will contribute to the determination of the country's position in the network. The sociogram allows to capture the multilateral effect on bilateral flows, giving to every country a position *relative* to all the other countries in the trade network and depending on the *entire* trading system. The benefit that we receive from representing

**Figure 3 –**  
The Network of World Trade in Goods (major two export partners,) 2007.



**Note:** For each country, only the export flows toward the first and second trade partner are considered. Country labels are the iso3 country codes. The size of the circle associated to each country is proportional to the number of inflows. Different colors correspond to different geographical regions. Trade data come from BACI-CEPII dataset. The network is drawn using Pajek. The sequence of Pajek commands necessary to reproduce the above figure is included in the Appendix.

international trade as a network of trade flows is therefore the possibility to visualize the effect of the relationship between the trading countries and the structure of the network itself, revealing patterns that are difficult to see using other approaches.

The network depicted in figure 3 is characterized by several features. Since we are accounting for just the two major export markets for every country, no specific weight is attached to the links, and the figure represents a directed unweighted (binary) network. By construction there is no disconnected component in the network (i.e. no country or group of countries is isolated from the rest of the network). As in figure 2 the size of the circle corresponding to a country is proportional to the number of receiving links, and is highly heterogeneous. In figure 3 highly connected nodes are generally placed at the center of the network (i.e. US, Germany, China and Japan (JAP), France (FRA) and the UK (BGR)), while less well connected countries are placed at the hedges of the figure. The structure of the network is both core-periphery and multipolar, with a leading role played by the main European economies (on the upper right) and the United States (on the bottom left). Japan (on the bottom centre) and the emerging economy of China hold a notable position in the network, acting as the third pole. Ancillary to the United States – and in some cases to China and other East and South Asian countries – is the position

of Latin American countries. On the other hand, Europe acts as an attractor for its close neighbors of North Africa, while central African countries look at Asia as a possible alternative. Former colonial links (or in general the role of history, e.g. common language, common institutional structure) emerges clearly from the picture: francophone African countries are close to France, while anglophone African countries are close to the UK. Also the role of physical distance (together with regional trade agreements) is evident: trade tends to be regionalized and countries sharing the same continental color tend to be close together. Interesting is the position of some Central and Eastern European countries, such as Latvia (LVA), Lithuania (LTU) and Estonia (EST) that hold a peripheral position but maintain preferential ties between them and with the nearest neighbors. It's worth nothing that – in spite of common perception – China and India hold a very different position both with respect to the US and in the network as a whole: although they both are important players, however China can boast a more central position compared to India (IND). Finally, the position of the United Kingdom is somehow peculiar, being located in between the US and the rest of its fellows European countries. Its position is however not right at the center of the network, depending not only on its direct links with the US and Germany, but also on being a major partner for Canada (CAN), Turkey (TUR), Norway (NOR), Kenya (KEN), Rwanda (RWA) and the Seychelles Islands (SYC). Those links are forcing the UK into a more decentralized position. On the contrary, countries such as Cambodia (KHM) and Bangladesh (BGD) are right at the center of the network, being linked only to the US and Germany. The same happens for Israel (ISR), the Dominican Republic (DOM), and Sierra Leone (SLE) being linked only to the US and Benelux (BLX). That means that the concept of centrality should not be automatically interpreted as a “good thing”. In fact, being central could reflect both an effective pivotal role or, conversely, a condition of strong dependence on major players and the absence of feasible alternatives.

To conclude, let's go back to the previous thought experiment, and let's delete again, now from figure 3, the link between Brazil and the US. Something totally different from the previous case will happen. Not only the circle corresponding to the US would become marginally smaller, but the position of many countries would be affected by the fact that Brazil is not exporting anymore to the US. Erasing the link with the US pushes Brazil away from the US and towards Argentina and Paraguay (PRY). If now we substitute the link between Brazil and the US with a new link between Brazil and Germany the whole Southern Cone would move towards Europe. Contrary to figure 2, this is the visual analog of the assumption of conditional interdependence among international trade flows.

From the next section on we will consider the entire World Trade Network and not only the two major export partners for each country. We will look at it as a directed network, both in its unweighted (i.e. binary) and weighted (trade flows) version.

## 2.1. Graphs and Networks: Basics

Let's define the world trade *network* as *graph* plus some extra-information. The graph  $\mathcal{G} = (\mathcal{V}, \mathcal{L})$  describes the set of dyadic relations among the vertices. Since we always work with the same number of countries, the  $\mathcal{V}$  dimension of the graph is  $\mathcal{V} = (1, \dots, 178)$ , and  $N=178$  is the total number of countries included in the analysis.<sup>12</sup> The number of links changes year by year: in 2007 the link-dimension of the graph is  $\mathcal{L} = (1, \dots, 22002)$ , where  $M$  is the total number of existing links. The links are directed (i.e. arcs), going from the exporting country,  $i$ , to the importing country,  $j$ , and  $\mathcal{L}_{ij} \in \{0, 1\}$ . The information included in  $\mathcal{L}$  is binary, and depicts the existence ( $\mathcal{L}_{ij} = 1$ ) or non-existence ( $\mathcal{L}_{ij} = 0$ ) of a trade link. If we add some more information on the link or the nodes, the graph becomes a network:  $\mathcal{N} = (\mathcal{G}(\mathcal{V}, \mathcal{L}), \mathcal{W}, \mathcal{P})$ . The *line value function*  $\mathcal{W}$  includes all relevant characteristics of the edges, while the *vertex value function*  $\mathcal{P}$  includes all relevant characteristics of the nodes. The  $\mathcal{W}_{ij}$  positive elements in  $\mathcal{W}$  act as dyadic weights on  $\mathcal{G}$ , transforming the graph in a directed weighted network: where  $\mathcal{W}_{ij}$  indicates the strength of the link between country  $i$  and country  $j$  (e.g. export volume). Instead,  $\mathcal{P}$  includes country-specific values (e.g. label, income, geographical coordinates). Some of the elements of the line and the vertex value functions are *exogenous*, like the ones indicated in parenthesis, others are *endogenously* determined from the topological properties of the network: the relative centrality of a node with respect to the structure of the network can be included in the vertex value function as any other characteristic of a node.<sup>13</sup>

The trade network is characterized by high dimensionality and strong heterogeneity. The 22002 existing links in 2007 correspond to a *density* of the network of 0.7 (i.e the ratio between the number of “realized” links and the number of maximum links possible), indicating that taking two countries at random the probability of an existing trade link among them is of 70%. It has evolved from around 50% in 1995 to the 68% of 2010, with a remarkable drop during the the big trade collapse, as shown in figure 4.

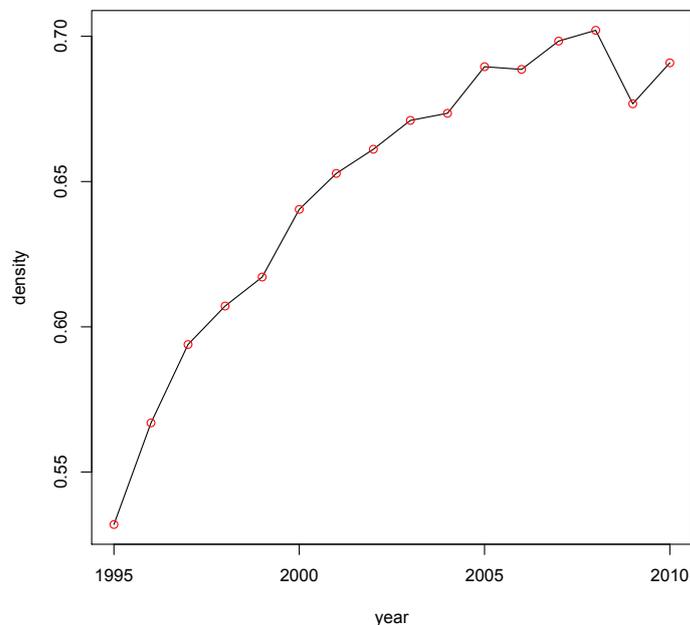
This matches with a very strong heterogeneity among trade flows: 25% of total world trade is made of just 32 links of higher strength  $\mathcal{W}_{ij}$ , involving only 17 countries; similarly, 50% of it is made of 158 links by 43 countries; 75% of it is made of 597 links by 76 countries; finally, 90% of total world trade is made of 1699 links (7.7% of total links) and 136 countries (76.4% of countries). More than 90% of bilateral trade flows are of modest relevance in their marginal contribution to world trade; at the same time one country out of four is equally marginal in its contribution to world trade.<sup>14</sup>

<sup>12</sup> The complete list of countries included in the analysis is in the Appendix, together with iso3 codes and geographical regions and geographical regions' colors.

<sup>13</sup> General overviews of the formal descriptions of graphs and networks can be found in Wasserman and Faust (1994), Newman (2010) and Jackson (2010). For an application to the World Trade Network see De Benedictis and Tajoli (2011).

<sup>14</sup> See section 2.2 on this, especially figure 6 and the comment below.

**Figure 4 –**  
Density, 1995-2010

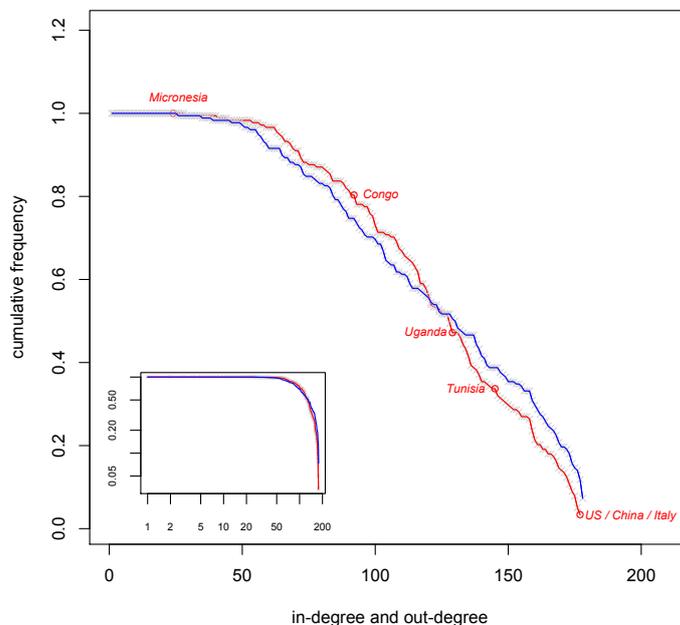


**Note:** Trade data come from BACI-CEPII dataset. The time-evolution of the density is drawn using R.

If the distribution of trade flows among country-pairs (i.e the weighted network) appears to be highly heterogeneous, similar heterogeneity is encountered also in trade partnerships (i.e. the unweighted network). Calling *out-degree* the number of out-going links originated from country  $i$  towards its trade partners (direct trade partners define the *nearest neighborhood* of  $i$ ), and *in-degree* the number of in-coming link to country  $i$ , it is possible to plot the cumulative distributions of such degrees, as in figure 5. The in-degree distribution is the one where some countries are identified by small circles: Italy, China and the US have the maximum possible number of in-degrees, importing from all possible countries in the database; Tunisia is importing from 132 countries; Congo from 92, and so forth. The out-degree distribution (with no countries identified) shows that the percentile of countries that trade with a high number of partners is higher for exports than for imports.

Both distributions are power-law (Clauset, Shalizi and Newman, 2009), and this is even more evident when the distributions are plotted on a log-log scale as in the framed inset in figure 5 (see also Riccaboni and Schiavo (2010) on this point). Even if most “laws” in economics are power-laws, this is a crucial finding. Extending an argument fully explored by Gabaix (2009, 2011), when the distribution of countries is fat-tailed, the central limit

**Figure 5** –  
Countries' Out-degree and In-degree in 2007: Cumulative Distribution



**Note:** The whole set of trade flows has been used in calculating the in-degree (the red continuous line with circles indicating some specific countries: US/China/Italy, etc.) and out-degree (the blue line) distributions. The inset show the same data on a log-log scale. Trade data come from BACI-CEPII dataset. The distribution is drawn using R (*igraph*).

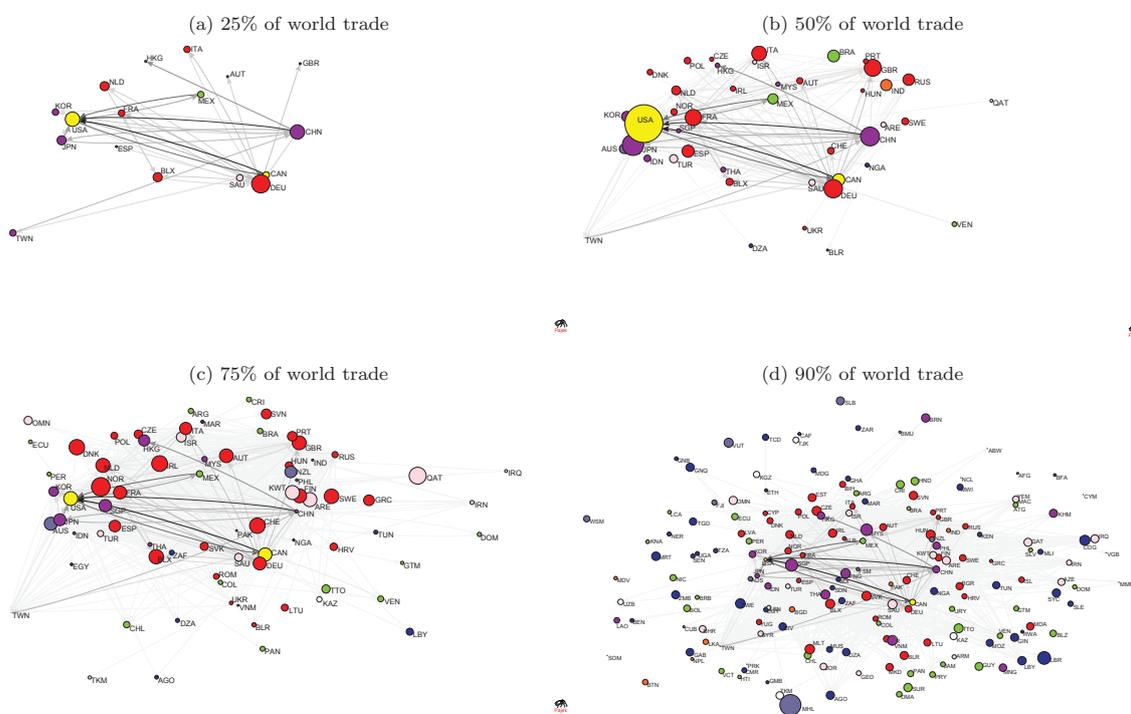
theorem breaks down, and idiosyncratic shocks to heavily connected countries explain a non-trivial fraction of aggregate world fluctuations.

## 2.2. Graphs and networks: sociograms

A further issue related to the high dimensionality of the World Trade Network regards the use of a sociogram as a visual tool. The possibility of drawing a network similar to the one in figure 3 including the full set of 22002 links is precluded by the excessive numerosity in the link-dimension. The resulting picture would not convey any clear information. In producing figure 3 we used a compression scheme on network tie strength, operating on the line value function  $\mathcal{W}$  and including for every country  $i$  only the two out-flows of higher strength (i.e. *out-strength*) with a weight  $\mathcal{W}_{ij} = 1$ , and imposing a weight  $\mathcal{W}_{ij} = 0$  on all other links. In figure 6 we opted for a different strategy in order to give emphasis to the strong heterogeneity in the distribution of the values of bilateral trade flows. As we previously mentioned, 25% of total world trade is accounted by the 32 links with the higher strength. In the first panel of figure 6, panel (a), will just plot a sociogram

including the major 32 links, and in subsequent panels (b), (c) and (d), we add further links in a descending order respect to the the strength of links. Furthermore, we choose to enrich the visualization of figure 6 adding some specific economic characteristics of single countries. To this aim, in each panel of figure 6, we make the nodes' dimension proportional to some endogenous attributes, such as the number of in-coming links (in panel (a)), or of some exogenous ones, such as GDP (in panel (b)), per capita GDP (in panel (c)); and openness (in panel (d)).

**Figure 6 –**  
The World Trade Network in Goods, 2007.



**Note:** Country labels are the iso3 country codes. Trade flows are represented on a gray scale, where light gray indicates bilateral trade flows of smaller entity and dark gray indicates major entities. The size of the circle associated to each country is proportional to the number of inflows (panel (a)), to GDP (panel (b)), to GDP per capita (panel (c)), and to the level of openness (panel (d)). GDP is in current US dollars in 2007. Different colors correspond to different geographical regions. Trade data come from BACI-CEPII dataset. The network is drawn using *Pajek*. The sequence of *Pajek* commands necessary to reproduce the above figure is included in the Appendix.

In figure 6, we zoom out, from panel(a) until panel (d), on the same set of data, to

emphasize the heterogeneity in the distribution of world trade flows in 2007. Panel (a) visualizes the top 32 bilateral trade flows, adding up to 25% of world exports. The strength of the flows is represented on a gray scale: exports from Canada (CAN) to the US (USA) rank first, followed by exports from China (CHN) to the same destination market, and so on. The world trade is dominated by a core group of 17 key players. Among them, the most important European countries, the North American countries, Japan (JAP) and some of the most dynamic East and Southeast Asian countries, Saudi Arabia (SAU), the only country from Middle East, and Mexico (MEX). To be noted that no other country from Latin America, besides Mexico, neither from Africa, is among the top world players. Furthermore, the presence of Taiwan (TWN) and Hong Kong (HKG) and the one of Benelux (BLX) and the Netherlands (NLD) is due to a worldwide “Rotterdam effect.”<sup>15</sup> The US, Germany, and China are the most interconnected countries in the world, as shown by the dark gray arcs between them. Remarkable is also the relationship among NAFTA members, the US, Canada, and Mexico.

In this panel the size of vertices is proportional to the number of outflows.<sup>16</sup> The UK (GBR) and Austria (AUT) receive only one trade inflow (from Germany (DEU)) with no edges indicating outflows; the two countries are therefore represented by small circles; Germany instead is sending trade outflows to many countries and is therefore represented by a big circle.

From the discussion of figure 3 we should have expected Hong Kong, Taiwan and China to be close together, and the same for Austria (AUT) and Germany. Taking the latter case as an example, we can see from panel (a) in figure 6 that Austria receives just one trade inflow, the one coming from Germany. This exclusivity should act as a unique attracting force, bringing the two countries close to each other. However, the position of the two countries, and of all other countries in all the panels in figure 6 depends not only on the links visualized in each panel, but on *whole* trade relationships included in the full dataset. This is why the position of countries in the *same* topological space represented in the subsequent four panels of figure 6 does not change. And this is why the position of Taiwan, or the one of the UK (GBR), looks peripheral in panel (a) of figure 6: the attracting forces of hidden nodes and arcs (that will become visible in subsequent panels) more than counterbalance the ones of the visible ones. This is again the visual result of countries’ interdependence.

If we increase the links up to 50% of world trade flows, the number of countries rises to 43,

<sup>15</sup> The “Rotterdam effect” is associated to the common practice of transshipment. Rotterdam in the Netherlands and Antwerp in Belgium are two of the biggest ports in the world, and handle substantial quantities of north European trade. Some of those imports are consumed in the Netherlands and Belgium, some others are shipped onwards to other European countries by road or rail; others get transferred to cargo vessels sailing to other continents.

<sup>16</sup> In section 3 we will call the number of inflows *in-degree* and the number of outflows *out-degree*. See Scott (2000) and Newman (2010) for a glossary of the terms used in network analysis.

generating 158 links. Figure 6 (panel (b)) shows this network emphasizing the economic dimension of countries, whose proxy variable is here total GDP (the size of vertices is now proportional to the country GDP in 2007). The GDP captures the so-called “size effect,” the well-known phenomenon whereby larger countries trade, in absolute terms, more than smaller ones (Helpman and Krugman, 1985).

The core countries are the largest economic countries as well, while a peripheral position is held by some small countries, belonging to different geographical regions (such as Qatar (QAT), Venezuela (VEN), Belarus (BLR), Ukraine (UKR), Algeria (DZA), and Taiwan (TWN)). Countries of middle size now appear as well: among them, some emerging economies, such as Brazil (BRA), India (IND) and Russia (RUS), located close to the UK; some North and East European countries, showing strong link with Germany; some East and South Asian economies, such as Indonesia (IDN), Singapore (SGP), Thailand (THA), and Malaysia (MYS), placed close to the US (except for the last one); and some Middle East countries, such as Israel (ISR), Turkey (TUR), the United Arab Emirates (ARE), and Qatar (QAT), holding very different positions in the network. Two African oil exporting countries belong to this club: Algeria (DZA) and Nigeria (NGA), although they keep a quite peripheral position.

Adding further links up to 75% of world trade, and characterizing countries by their per capita GDP (as showed by the size of nodes) to capture the so-called “income effect” (i.e. richer countries trade more than poorer ones), we get a larger network, which now comprises 76 countries (figure 6 (panel (c))). The countries of the core structure are almost all rich countries as well. This panel emphasizes the emerging of a peripheral shape with the inclusion of several low income countries (excluding the Arabian Peninsula countries, such as Qatar (QAT) and Oman (OMN)). It’s worth nothing the appearance of many Latin American countries as well as few African countries, although placed in a semi peripheral position (with the exception of South Africa (ZAF) placed in the central part of the network).

Lastly, extending the links up to 90% of total world trade (i.e. including flows of modest relevance), the network consists of 136 countries. In figure 6 (panel (d)), the size of the circle corresponding to a country is proportional to its openness (i.e. export to GDP). Here the network shows a definite core-periphery shape. Furthermore, it emphasizes a well-known notion of the trade literature, that is small countries tend to have more open economies than large countries. It is particularly apparent for countries belonging to the East and South East Asia, placed in a central position, such as Singapore (SGP) and Malaysia (MYS), as well as remote regions such as Marshall Islands (MHL), Vanuatu (VUT), the Solomon Islands (SLB), and Brunei Darussalam (BRN) or some African countries, such as Liberia (LBR) and Congo (COG), all located at the border of the network.

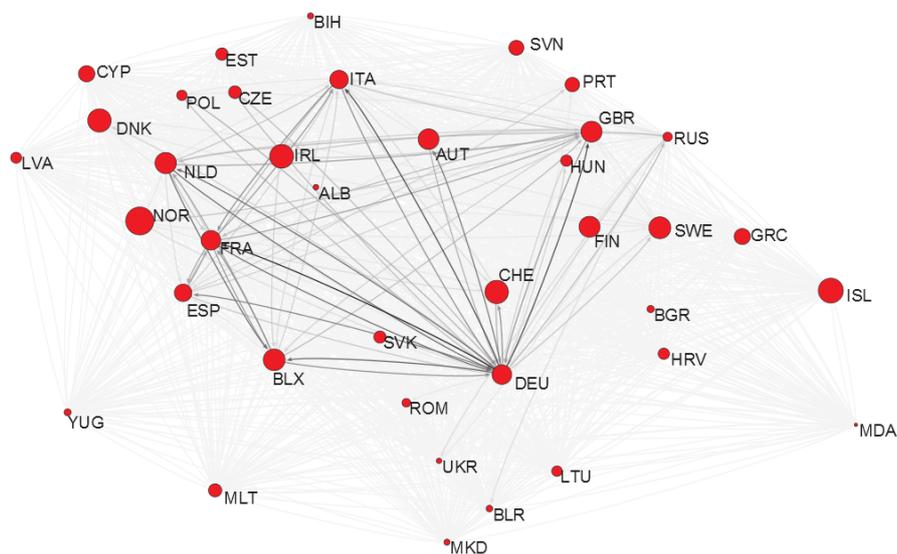
### 2.3. Zooming in and out trade data: Local and Global views

If we are interested in inspecting the structure of trade ties not at global level but in a specific part of the network, for example a continent, we could zoom in the network on countries belonging to the same continent. Figure 7 presents trade links within Europe, where the size of the vertices is proportional to the value of per capita GDP.

This visualization permits to better identify the position of each European country within the overall trade network. The structure of this regional network has a polyhedral and a clear core-periphery form. The main economies of the European Union are the core group ( Benelux (BLX), France (FRA), Germany (DEU), Italy (ITA), the Netherlands (NLD), and the United Kingdom (GBR)). Six of them are the founding members of the European Economic Community, and are members of the Economic and Monetary Union as well. A peripheral position is displayed by Malta (MLT) and the current EU candidates (Iceland (ISL) and those belonging to the former Yugoslavia, such as Serbia and Montenegro (YUG), Macedonia (MKD) and Croatia, (HRV)). Also Moldova (MDA) seems to be weakly connected to the European network, holding the most peripheral position among the countries of the Community of Independent States. The “income effect” is here less evident. Many high income countries are not in the center of the network, located on the right part (Iceland (ISL), Finland (FIN), and Sweden (SWE)) or on the left part (Denmark (DNK) and Norway (NOR)) of the network. It is plain also to detect the central role played by Germany (DEU): it is the country with the strongest links (observable through the dark gray lines) with the most important European partners. Albania (ALB), notwithstanding its low income level, is strongly connected to the core countries. As above underlined, this kind of position denounces a condition of strong dependence on major players rather than a key role in the network. The same goes also in the case of Ireland (IRL) and Austria (AUT). It is worth noting the position of Switzerland (CHE) and Slovakia (SVK) largely connected to the network and with a strong tie with Germany. Quite peculiar the position of Poland (POL) and Hungary (HUN) curiously away one from the other, with the first one holding a less central position but strongly connected with Germany.

If our interest is instead to analyze the relations between aggregate geographical regions to detect, for instance, which continents or sub-continents have strong trade ties, we may zoom out to obtain a global view of the network (see figure 8). All exports of European countries to North American countries, for instance, are now replaced by one new edge pointing from North America to European countries. The new line indicates the value of exports that is equal to the sum of all original trade values. The global view shows now the position of Europe into the world system. Darker gray lines indicate stronger trade links between geographical regions, while exports within a geographical region are displayed by loops (lines that connect a vertex to itself). From figure 8 it is plain to detect the key role played by Europe in the world trade network, as underlined by both the strong ties

**Figure 7 –**  
Trade links within Europe, 2007.



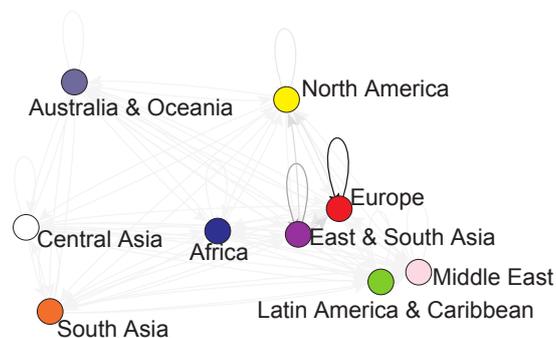
**Note:** Country labels are the iso3 country codes. The size of the circle associated to each country is proportional to the value of GDP per capita (GDP per capita in current US dollars in year 2007). Trade flows are described using a gray-scale, from light (small flows) to dark (large flows). Trade data come from BACI-CEPII dataset. The network is drawn using Pajek. The sequence of Pajek commands necessary to reproduce the above figure is included in the Appendix.

with North America and East and South East Asia and the dark loop, which reveals the strongest intra-area trade exchange among all the geographical regions. As relevant the place of North America and East and South East Asia in the network, while quite peculiar is the position of Latina America and Caribbean away from North America and closer to Europe and East and South Asia. It is interesting to notice the central position holds by Africa in this visualization: Africa accounts for a small percentage of world trade (as described in figure 3) and this position is mainly due to its high dependence on (i.e. the attracting force of) European and East and South Asian players.

### 3. Local and global centrality measures and relative countries' position

In section 2.1, we have visualized the “structural” characteristics of the trade network. Now we want to focus on “ego” measures, i.e. how and how much each single country (i.e. node) is relatively positioned in the overall network, considering the trade relations with *all* countries inside and outside the geographical region to which it belongs. Country heterogeneity will be highlighted in terms of the position of each country in the trade

**Figure 8 –**  
Trade among geographical regions, 2007.



**Note:** Country labels are the iso3 country codes. The size of the circle associated to each country is proportional to the value of GDP per capita (GDP per capita in current US dollars in year 2007). Different colors correspond to different geographical regions. Trade flows are described using a gray-scale, from light (small flows) to dark (large flows). Trade data come from BACI-CEPII dataset. The network is drawn using *Pajek*. The sequence of *Pajek* commands necessary to reproduce the above figure is included in the Appendix.

network, and will allow to endogenously rank countries according to their relative *centrality*. Network analysis provides several indicators to assess the importance of a node centrality, capturing different aspects of its position (Borgatti, 2005). In general, each of them can provide different types of information, the usefulness of which depends on which kind of relationship the network under analysis is describing (information flows, parental relationship, bargaining power, infection transmission, etc.). It follows that any particular measure will be better suited for some applications and less well suited for others. Here, we give a brief review of these indicators, concentrating our attention on the ones that we think are more suitable to the description of the World Trade Network.

### 3.1. Centrality measures

Centrality measures can be classified into four main groups (Jackson, 2010): a) degree centrality,  $C_D$ , that measures how a node is connected to others (with strength centrality

$C_S$  as a weighted version of  $C_D$ ); b) closeness centrality,  $C_C$ , showing how easily a node can be reached by other nodes; c) betweenness centrality, describing how important a node is in terms of connecting other nodes; d) the fourth group of indexes, such as the eigenvector centrality measure,  $C_E$ , or the Bonacich centrality,  $C_B$ , associates node's centrality to the node neighbors' characteristics, directly referring to how important, central, influential or tightly clustered a node's neighbors are. We focus on three of them, limiting the description of the betweenness centrality to a footnote, since we consider it less suitable for the trade data used in this analysis.

### 3.1.1. Degree centrality

*Degree centrality* is the simplest measure of the position of a node in a network. If the network is unweighted, it measures the centrality of a node by the number of connections the node has.

$$C_D = d_i(\mathcal{L}) = \#\{i : \mathcal{L}_{ij} = 1\} = \sum_{j \neq i}^N \mathcal{L}_{ij}. \quad (1)$$

Let's recall that  $N$  is the total number of nodes (countries) in the network, and  $\mathcal{L}_{ij}$  is the element  $(i, j)$  in the trade adjacency matrix  $\mathcal{L}$ , where  $i$  is the row-indicator corresponding to exporting countries, and  $j$  is the column-indicator corresponding to importing countries. If  $\mathcal{L}_{ij} = 1$  the two countries  $i$  and  $j$  are trade partners (regardless of the direction of trade flow, i.e. the network is undirected), if  $\mathcal{L}_{ij} = 0$  they are not trading between each others.

The degree centrality measure, being dependent on the number of the existing nodes in the network, makes it difficult to compare networks of different node-size. Even if this is not what we have in the present version of the World Trade Network (being the number of countries included in the analysis constant) it is usually better to calculate the normalized version of  $C_D$ , using the total number of possible neighbors excluding self,  $N - 1$ , as a normalized factor:

$$C_D^N = \frac{\sum_{j \neq i}^N \mathcal{L}_{ij}}{(N - 1)}. \quad (2)$$

It follows that this indicator ranges from 0 to 1; the more is the degree centrality close to 1, the more a country is directly connected to the rest of the network.

In a directed network there will be two measures of degree centrality: *in-degree* centrality, measuring the number of arcs pointing to ego, and *out-degree* centrality, measuring the number of arcs exiting from ego.<sup>17</sup> In the case of out-degree,  $\sum_{j \neq i}^N \mathcal{L}_{ij}$  is the total number

<sup>17</sup> The sociological literature associates the terms *prestige* to in-degrees and *influence* to out-degrees.

of countries toward which country  $i$  is exporting; in the case of in-degree,  $\sum_{j \neq i}^N \mathcal{L}_{ji}$  is the total number of countries from which country  $i$  is importing,

$$C_{D_{out}}^N = \frac{\sum_{j \neq i}^N \mathcal{L}_{ij}}{(N-1)} \quad C_{D_{in}}^N = \frac{\sum_{j \neq i}^N \mathcal{L}_{ji}}{(N-1)}. \quad (3)$$

The degree centralities  $C_{D_{out}}^N$  and  $C_{D_{in}}^N$ , defined in equation 3, are summarized in columns (1) and (2) in table 1.<sup>18</sup> Aruba, the first country in alphabetical order in our dataset, has an in-degree centrality of 0.37 and an out-degree centrality of 0.36. Generally, the two measure are not much different one from the other, but exceptions are possible. Armenia has an in-degree centrality of 0.73 and an out-degree centrality of 0.58. The country is receiving links from 73% of the countries in the network, while is sending links to 58% of them. The ratio of the two measures indicates the relative connectivity of a country in terms of inflows (with a ratio greater than 1) or outflows (with a ratio smaller than 1). Armenia, with a degree-ratio of 1.26 is better structurally connected in terms of in-degree than in out-degree.

If we normalize the number of links in equation 1 by the total number on links  $M$  in the network, we obtain the following measures:

$$C_{D_{out}}^P = \frac{\sum_{j \neq i}^N \mathcal{L}_{ij}}{M} \quad C_{D_{in}}^P = \frac{\sum_{j \neq i}^N \mathcal{L}_{ji}}{M}. \quad (4)$$

In this case, the out-degree centrality measure  $C_{D_{out}}^P$  indicates the number of country's out-degree in percentage terms, and similarly for  $C_{D_{in}}^P$ . The degree centralities defined in equation 4 are summarized in columns (3) and (4) in table 1. Aruba's degree centralities in percentage terms are  $C_{D_{in}}^P=0.30$  and  $C_{D_{out}}^P=0.29$ , meaning that among the  $M=22002$  links in the 2007 trade network, Aruba is responsible of the 0.30% of unweighted inflows and 0.29% of unweighted outflows.

Moving to the weighted version of the network, we now examine *strength centrality*, considering trade volumes instead of trade partnerships. Given the very strong heterogeneity

<sup>18</sup> We calculated the centrality measures in equations 3 and 5 for the year 2007 using the Stata Graph Library (SGL). The full panel 1995-2010 of (these and subsequent) centrality measures for 178 countries can be downloaded from the [CEPII databases page](#). The structure of the downloadable data is the same of tables 1 and 2. In table 1 (for the unweighted trade network) and table 2 (for the weighted trade network) we include, for expositional purposes, just a subset of the countries in the sample, the first and last ten countries in alphabetical order. We discuss and compare the different outcomes obtained using different centrality measures in the unweighted and weighted trade network in section 3.2.

Table 1 – Centrality - unweighted measures

id	i	iso3	country	(1) Out-degree	(2) In-degree	(3) Out-degree percent	(4) In-degree percent	(5) Out-closeness	(6) In-closeness	(7) Out-eigenvector	(8) In-eigenvector
1	533	ABW	Aruba	.35593	.37288	.28634	.29997	.60825	.61458	.0396	.04252
2	4	AFG	Afghanistan	.49718	.48588	.39996	.39087	.66541	.66045	.05651	.05558
3	24	AGO	Angola	.42373	.55367	.34088	.44541	.63441	.69141	.04887	.06264
4	8	ALB	Albania	.62712	.71186	.5045	.57268	.7284	.77632	.06842	.07549
5	530	ANT	Netherlands Antilles	.53107	.46328	.42723	.37269	.68077	.65074	.05777	.05198
6	784	ARE	United Arab Emirates	.9322	.89831	.74993	.72266	.93651	.90769	.09198	.09089
7	32	ARG	Argentina	.93785	.76271	.75448	.61358	.94149	.80822	.09273	.07971
8	51	ARM	Armenia	.57627	.73446	.46359	.59086	.70238	.79018	.06265	.07706
9	28	ATG	Antigua and Barbuda	.51977	.64972	.41814	.52268	.67557	.74059	.05571	.06862
10	36	AUS	Australia	.9887	.9661	.79538	.7772	.98883	.96721	.09499	.09479
...	...	...	...	...	...	...	...	...	...	...	...
169	92	VGB	Virgin Islands	.49153	.44633	.39542	.35906	.66292	.64364	.05309	.04982
170	704	VNM	Viet nam	.9209	.77966	.74084	.62722	.9267	.81944	.09161	.08121
171	548	VUT	Vanuatu	.31638	.36158	.25452	.29088	.59396	.61034	.03615	.04019
172	882	WSM	Samoa	.24859	.32203	.19988	.25907	.57097	.59596	.02909	.03725
173	887	YEM	Yemen	.59887	.66102	.48177	.53177	.71371	.74684	.06487	.07087
174	891	YUG	Serbia and Montenegro	.83616	.88701	.67267	.71357	.85922	.89848	.08553	.08965
175	711	ZAF	South Africa	.9887	.98305	.79538	.79084	.98883	.98333	.09528	.096
176	180	ZAR	Congo Dem. Rep.	.40678	.41243	.32724	.33179	.62766	.62989	.04611	.04724
177	894	ZMB	Zambia	.55932	.66667	.44996	.53631	.69412	.75	.06171	.07135
178	716	ZWE	Zimbabwe	.77966	.59887	.62722	.48177	.81944	.71371	.07986	.06467

Table 2 – Centrality - weighted measures

id	i	iso3	country	(1) Out-strength	(2) In-strength	(3) Out-strength percent	(4) In-strength percent	(5) W-Out-closeness	(6) W-In-closeness	(7) W-Out-eigenvector	(8) W-In-eigenvector
1	533	ABW	Aruba	24328.21	14732.21	.03207	.01942	.72634	.35986	.00257	.00142
2	4	AFG	Afghanistan	1880.917	20843.3	.00248	.02747	.11901	.31131	.00003	.00108
3	24	AGO	Angola	187955	68658.4	.24775	.0905	.82343	.37339	.02293	.00396
4	8	ALB	Albania	6845.41	24469.3	.00902	.03225	.48718	.34375	.00032	.00093
5	530	ANT	Netherlands Antilles	19876.57	26701.99	.0262	.0352	.46844	.37334	.00089	.00303
6	784	ARE	United Arab Emirates	788635	694503.6	1.03952	.91544	.84608	.41277	.04276	.04072
7	32	ARG	Argentina	323668.9	244444.3	.42664	.32221	.80289	.40881	.01575	.01512
8	51	ARM	Armenia	7777.557	18172.09	.01025	.02395	.23745	.28337	.00031	.0007
9	28	ATG	Antigua and Barbuda	1490.491	9677.804	.00196	.01276	.12058	.25492	.00006	.00076
10	36	AUS	Australia	823285.5	869600	1.08519	1.14624	.84254	.41627	.06398	.06164
...	...	...	...	...	...	...	...	...	...	...	...
169	92	VGB	Virgin Islands	2261.846	21634.1	.00298	.02852	.17193	.36645	.0001	.00076
170	704	VNM	Vietnam	307967.3	361909.4	.40594	.47704	.82077	.41135	.02439	.0185
171	548	VUT	Vanuatu	1990.4	1511.497	.00262	.00199	.20592	.08091	.00012	.00006
172	882	WSM	Samoa	989.1105	1631.536	.0013	.00215	.1976	.09984	.00003	.00007
173	887	YEM	Yemen	40926.56	60153.24	.05395	.07929	.63469	.34658	.0028	.00314
174	891	YUG	Serbia and Montenegro	50779.54	117439.3	.06693	.1548	.6137	.37973	.00158	.00451
175	711	ZAF	South Africa	502145.9	460072.9	.66189	.60643	.81734	.40969	.03207	.02632
176	180	ZAR	Congo Dem. Rep.	11157.5	15483.54	.01471	.02041	.41192	.29855	.00084	.00055
177	894	ZMB	Zambia	34256.3	22570.01	.04515	.02975	.67041	.36212	.00138	.00061
178	716	ZWE	Zimbabwe	21609.34	18506.23	.02848	.02439	.6039	.35782	.00074	.00051

among trade flows, as reported in section 2.1, strength centrality measures will differ from degree centralities. Unweighted and weighted networks conceive different information (Newman, 2010).

In equation 5, the centrality measure are computed aggregating over of the weights of the arcs (export or imports flows) connected to the node and normalizing by  $(N - 1)$ . The index we obtain is a measure of the *out-strength* and *in-strength*.

$$C_{S_{out}}^N = \frac{\sum_{j \neq i}^N \mathcal{W}_{ij}}{(N-1)} \quad C_{S_{in}}^N = \frac{\sum_{j \neq i}^N \mathcal{W}_{ji}}{(N-1)}. \quad (5)$$

The strength centralities defined in equation 5 are summarized in columns (1) and (2) in table 2. Aruba has an ‘average’ flow of imports of 14732 thousands of US dollars, and an ‘average’ flow of exports of 24328 thousands of US dollars.<sup>19</sup> The prevalence of  $C_{S_{out}}^N$  over  $C_{S_{in}}^N$  indicates an overall trade surplus.

If we normalize the in-strength and out-strength by total world trade ( $\sum_i \sum_j \mathcal{W}_{ij} = \sum_i \sum_j \mathcal{W}_{ji}$ ) we obtain:

$$C_{S_{out}}^P = \frac{\sum_{j \neq i}^N \mathcal{W}_{ij}}{\sum_i \sum_j \mathcal{W}_{ij}} \quad C_{S_{in}}^P = \frac{\sum_{j \neq i}^N \mathcal{W}_{ji}}{\sum_i \sum_j \mathcal{W}_{ji}}, \quad (6)$$

which measure the strength centralities in percentage terms. Obviously, the measures defined in equation 6, and summarized in column (3) and (4) in table 2, are plain vanilla trade shares. Aruba has a market share of 0.0194% over total exports and a share of 0.0321% over total imports. Switching to big players, in 2007 China had an export share of 10.2% and an import share of 5.5%, while the US had an export share of 8.3% and an import share of 14.0%.

As you can see, some network centrality measures are closely related or are perfectly matching more traditional trade statistics, such trade shares. Therefore, what’s the point of describing and discussing them? We will come back to this point later, but as a prelim of a deeper discussion it is sufficient to say that in the next sections we will compare this network measures with other more complex measures as a way to relate more traditional trade statistics with network centrality measures that take the whole structure of the network into account. This as a way to highlight the difference between networks centrality measure and alternative trade indicators.

To conclude this presentation of degree and strength centrality measures it is noteworthy that weighted measures should not be considered as an improvement to unweighted measures. Both measures point to different aspects. Degree centrality puts the emphasis on the number of trade links disregarding the specific weights of the link. As can be seen in tables 1 and 2, Australia and South Africa share the same out-degree centrality ( $C_{D_{out}}^N = 0.99$ ) while the out-strength centrality points out that Australia has an export

<sup>19</sup> The measure is an ‘average’ with respect to the possible maximum number of trade partners, not the realize number of trade partners. Dividing  $C_S^N$  by  $C_D^N$  generates a proper algebraic mean of aggregated flows of imports and exports, respectively.

share  $C_{S_{out}}^N$  almost double that of South Africa. On the other hand, weighted measures, putting emphasis on trade volumes, disregard the numerosity of trade partnerships of a country, and countries with few links of high strength or many links of lesser strength will be accounted for in an equal way. Unweighted and weighted measures should be looked in parallel, exploiting the joint information coming from the two of them.

Those unweighted and weighted measures have an obvious counterpart in the notion of trade partners and trade volumes (Helpman, Melitz and Rubinstein, 2008), and can also be manipulated to match naturally with the notion of intensive an extensive margin of trade (Felbermayr and Koheler, 2005). Since in our case the number of nodes is fixed, a variation in the weights of already existing trade links in the weighted trade network is equivalent to the notion of intensive margin, while the activation of previously non-existing bilateral trade flows (i.e. new weighted links) is equivalent to the notion of extensive margin.

The degree and the strength centrality are essentially *local centrality measures*. They take into consideration only the *direct links* of a node, its *nearest neighborhood*, regardless to the position of the node in the network's structure. In the network represented in figure 3 all countries have by construction the same out-degree, i.e.  $C_{D_{out}}^N=2$ , however, they attain very different positions in the network. Analogously, Paraguay (PGY) and Egypt (EGY) have both the same out-degree, as all the other countries have, *and* the same in-degree  $C_{D_{in}}^N=0$ ; however, Paraguay is at the south-periphery of figure 3, while Egypt is at center of the figure. Their positions depend in fact on *the position of other countries*: Paraguay is linked with peripheral countries, while Egypt is linked with very central countries. To uncover the effect of others we should move from local centrality measures to *global centrality measures*.

### 3.1.2. Closeness centrality

One of the most commonly used *global centrality measures* is the *closeness centrality*. It is a measure of how close (in terms of topological distance) a node is with respect to all other nodes. In general terms, the concept of distance in network analysis is related to the number of steps needed for some node 'to reach' another node in the network. The shortest path between country  $i$  and country  $j$  is called the *geodesic distance* between  $i$  and  $j$ .

Taking the inverse of the average geodesic distance as a measure of proximity, closeness centrality provides high centrality scores to nodes that are located closer to the set of reachable nodes:

$$C_C^N = \frac{(N-1)}{\sum_{j \neq i}^N \mathcal{D}_{ij}}, \quad (7)$$

where  $\mathcal{D}_{ij}$  is the number of steps in the shortest path between  $i$  and  $j$ . If country  $i$  is directly linked to all the other countries, the variable  $\mathcal{D}_{ij}$  would be 1 for every country-pair connection. In this case the denominator of equation 7 would be equal to  $(N-1)$  and  $C_C^N=1$ . The lesser one country is directly linked to others in terms of bilateral trade flows, the lower will be its closeness centrality.

Since the numerator of equation 7 is also the minimum possible value reachable by the denominator, the closeness centrality can also be interpreted as the inverse of the proportion between the shortest of the shortest paths and the sum of the actual shortest paths. In other terms, it is a distance measure between the actual topological distance and the minimum possible one.

As for the case of the degree centrality, we must consider the directionality of trade flows.

$$C_{C_{out}}^N = \frac{(N-1)}{\sum_{j \neq i}^N \mathcal{D}_{ij}} \quad C_{C_{in}}^N = \frac{(N-1)}{\sum_{j \neq i}^N \mathcal{D}_{ji}}. \quad (8)$$

In this case, the out-closeness centrality  $C_{C_{out}}^N$  indicates the sum of the geodesic distances of country  $i$ , normalized by the maximum number of possible export partners.  $C_{C_{in}}^N$  indicates a similar measure for import partners. The closeness centralities defined in equation 8 are summarized in columns (5) and (6) in table 1 and discussed in section 3.2. Continuing the dissection of the Aruba case, the out-closeness centrality is equal to  $C_{C_{out}}^N=0.61$ , which results from the inverse of the sum of Aruba's geodesic distances ( $\sum_{j \neq i}^N \mathcal{D}_{ij}=290$ ), normalized by  $(N-1)=177$ . Counting 66 export partners for Aruba, its average geodesic distance is 4.39, i.e. Aruba is on average 4.39 topological steps away from all other countries in the net. If we now consider Australia or South Africa – two cases of well connected countries exporting to 175 foreign countries and having an out-degree centrality very close to 1, i.e.  $C_{D_{out}}^N=0.99$  – their out-closeness centrality  $C_{C_{out}}^N$  is quite similar to their out-degree centrality. In those cases, when the sum of the geodesic paths approximates the degree of a node, *global centrality measures* are equivalent to *local centrality measures*. On the contrary, the lesser a country is directly connected to other countries, the more the two measures will separate from each others. We will further discuss this issue in section 3.2.<sup>20</sup>

<sup>20</sup> The same happens for other global measures, such as the *betweenness centrality*. This measure of centrality depicts how well situated a node is in terms of the path that it lies on. It assign highest score to nodes that lie on a larger proportion of shortest paths linking pairs of other nodes. Betweenness centrality is a useful measure in the cases when a node is important as an intermediary i.e. in a communication network. A node is crucial in the transmission of information if flows of information are disrupted or must make longer path if that node stops passing information or if it disappear from the network. However, in the case of aggregate or sectoral trade flows, this measure seems not well suited for an application to bilateral trade relationship, unless if it would be possible to distinguish among trade flows in final products and intermediate products.

In the case of closeness centrality, moving from the unweighted to the weighted version of the centrality is less straight forward than with degree centrality. A weighted path is in fact a mixture of two components: the length of the path (i.e. the minimum number of steps between  $i$  and  $j$ ) and the strength of the single steps of the path (i.e. bilateral trade flows). If one takes the components in a multiplicative way, the second component would inflate the weighted length of a path between two countries, *reducing* the centrality of the relevant country. This is what the SGL routine and the `igraph` library do by default. We will therefore follow a procedure proposed by Newman (2010) and Opsahl (2009) and implemented in the `tnet` R library.

We started from the algorithm elaborated by Dijkstra (1959) and apply it to a transformation of the original trade weighted matrix. In this case the elements in the line value function are not anymore the bilateral trade volumes  $\mathcal{W}_{ij}$ , but the share between  $\mathcal{W}_{ij}$  and the average bilateral trade volume in world trade:  $\omega_{ij} = N \frac{\mathcal{W}_{ij}}{\sum_i \sum_j \mathcal{W}_{ij}}$ . In this case, the Dijkstra (1959) algorithm calculates the weighted geodesic distance over  $\omega_{ij}$ :

$$l_{ij} = \min \left( \frac{1}{\omega_{iz_1}} + \frac{1}{\omega_{iz_2}} + \dots + \frac{1}{\omega_{z_{n-3}j}} + \frac{1}{\omega_{z_{n-2}j}} \right), \quad (9)$$

where the  $z$ s are the intermediate steps necessary to reach  $j$  from  $i$ . The resulting normalized *weighted closeness centrality* for directed networks is

$$C_{C_{out}}^{NW} = \frac{(N-1)}{\sum_{j \neq i}^N l_{ij}} \quad C_{C_{in}}^{NW} = \frac{(N-1)}{\sum_{j \neq i}^N l_{ji}}. \quad (10)$$

The weighted closeness centralities defined in equation 10 are summarized in column (5) and (6) in table 2 and discussed in section 3.2. Aruba has a  $C_{C_{out}}^{NW}=0.726$  and a  $C_{C_{in}}^{NW}=0.359$ ; Australia has a  $C_{C_{out}}^{NW}=0.842$  and instead a  $C_{C_{in}}^{NW}=0.416$ . Comparing the unweighted and the weighted closeness centralities, the distance between Australia and Aruba has decreased in the weighted version of the measure.

What does  $C_{C_{out}}^{NW}=0.842$  mean? The index can be interpreted as the inverse of the average weighted geodesic distance from  $i$  to its  $(N-1)$  potential trade partners.<sup>21</sup> In the Australian case, this is  $1.187 = \frac{1}{0.842}$ , having  $\omega_{ij}$  as the metric, i.e. Australia is 1.187 steps away from the rest of world countries, where steps are measured in terms of the average bilateral trade flow in world trade. Aruba is 1.377 steps away; Antigua and Barbuda are 8.333 steps away on average.

<sup>21</sup> When country  $i$  is not directly connected to country  $j$ , the  $l_{ij}$  is imputed using the average of the shortest weighted path from  $i$  to  $j$  (Dijkstra, 1959).

In both the Australian and the Aruban cases  $C_{out}^{NW} > C_{in}^{NW}$ . In the weighted version of the index, the prevalence of the out-closeness over the in-closeness is verified in the large majority of the cases. In the unweighted version of it, the positive sign in the difference between out-closeness and in-closeness is less prevalent than in the weighted version. Moreover, the sign is not invariant moving from the unweighted to the weighted version of the closeness centrality measure, and only the countries where  $C_{out}^N \gg C_{in}^N$  tend to have a  $C_{out}^{NW} > C_{in}^{NW}$ . The opposite is however not true.

As for the the strength-centrality measure, the weighted closeness centrality does not put emphasis on the number of links in the paths considered. The two measures, the unweighted and the weighted closeness centrality should be considered in a complementary way.

### 3.1.3. Eigenvector centrality

In addition to closeness (and betweenness) centrality, we can describe another and different perspective on global centrality. If closeness was stressing the relevance of links, here we stress the relevance of nodes. As we showed in figure 3, it is not irrelevant if a node is connected to central players or to peripheral ones. The basic idea is that a node's *eigenvector centrality* is determined by the eigenvector centrality of its neighbors. It is not the country's centrality itself that matters, what really matters is the centrality of the countries linked to him. The circularity of the argument is evident, but can be tackled using some matrix algebra.

Starting from a binary trade-matrix, it is possible to use the Bonacich (1972) eigenvector centrality to define the centrality of country  $i$  as the sum of the eigenvector centralities of its neighbors. That is:

$$C_E(i) = \mathcal{L}_{i1}C_E(1) + \mathcal{L}_{i2}C_E(2) + \dots + \mathcal{L}_{i(n-1)}C_E(n-1) + \mathcal{L}_{in}C_E(n). \quad (11)$$

The system of equations, including the eigenvector centrality of all  $n$  countries, can be rewritten in matrix form as:

$$(I - \mathcal{L})\vec{C}_E = 0, \quad (12)$$

where  $I$  is a  $n \times n$  identity matrix,  $\mathcal{L}$  is the trade adjacency matrix introduced in section 2.1, and  $\vec{C}_E$  is the  $n \times 1$  vector of countries' eigenvector centralities. Equation 12 is the

trade-matrix characteristic equation for an eigenvalue  $\lambda=1$ .<sup>22</sup> From the Perron-Frobenius theorem we know that a square matrix with positive (and some classes of nonnegative) real entries has a unique largest real eigenvalue and that the corresponding eigenvector has strictly positive components. The normalization we choose, as well as for a row-normalized or a column-normalized trade-matrix,<sup>23</sup> allows to exploit this property and to consider the entries of the relevant eigenvector as a measure of country centrality.

The eigenvector centralities defined in equation 12 are summarized in columns (7) and (8) in table 1 and discussed in section 3.2. Aruba has an out-eigenvector centrality equal to  $C_{E_{out}}^N=0.039$ , and an in-eigenvector centrality equal to  $C_{E_{in}}^N=0.042$ . Australia has an out-eigenvector centrality equal to  $C_{E_{out}}^N=0.094$ , and an in-eigenvector centrality equal to  $C_{E_{in}}^N=0.094$ . In general, countries with a high value of eigenvector centrality are the ones which are connected to many other countries which are, in turn, connected to many others. The largest values correspond to countries in large and cohesive (high-density) sub-networks. In spite of being a measure of *global centrality* the eigenvector centrality is highly correlated with the degree centrality.<sup>24</sup> A fortiori, given the high density of the World Trade Network.

Some problems may arise in case of a poorly connected graph (Newman, 2010). In case of a sharp segmentation of the network (e.g. extreme trade regionalization with disconnected regional components)<sup>25</sup> *all* countries would have an eigenvector centrality of zero, as a consequence of the circularity in the system of equations 11. To avoid this shortcoming, a variation of equation 12 has been proposed by Bonacich (1987). Here the *Bonacich centrality*  $C_B$  is defined as:

$$C_B = (I - \beta \mathcal{L})^{-1} \alpha \mathcal{L} u, \quad (13)$$

<sup>22</sup> Bonacich (1987) proposes also a variant of equation 12 were  $\lambda \neq 1$ . In this case,  $(\lambda I - \mathcal{L}) \vec{C}_E = 0$ , and the centrality of country  $i$  would be *proportional* to the centralities of its trade partners. Negative values of  $\lambda$  indicate conflictual relations (e.g. zero sum games on market shares), positive values indicate mutual gains.

<sup>23</sup> This is exactly the procedure proposed by Katz (1953). In the case of the *prestige centrality index* that takes its name, he proposed to adopt a column normalized matrix where the entries were the ratio  $\frac{\mathcal{L}_{ij}}{d_j(\mathcal{L})}$ . The relevant matrix becomes a column stochastic matrix. Extending Katz's procedure to a row normalized matrix an analogous centrality index. Often the names of Bonacich and Katz are associated in the denomination of this kind of eigenvector centrality measures. See also Bonacich and Lloyd (2001) and Newman (2010) on this issue. In `SGL` and `igraph` the standard option operates on the transpose of  $\mathcal{L}$  and the in-eigenvector centrality is produced by default (Miura, 2012) and (Csardi and Nepusz, 2006).

<sup>24</sup> This is not a case: the first eigenvector of  $\mathcal{L}$  is closely related to the rank-one decomposition of the adjacency matrix.

<sup>25</sup> In this case the trade matrix would be characterized by block-diagonality and the networks would show some disconnected components.

where  $\mathbf{u}$  is a  $1 \times n$  vector of ones, and  $\alpha$  and  $\beta$  are scalars, and  $\beta$  is sufficiently small ( $\beta < 1/\lambda$ ) so that equation 13 is well defined. In this formula, the additional inclusion of  $\beta$  (as a decay parameter) makes the  $C_B$  of  $i$  proportional to the centrality of other countries, while  $\alpha$  (as a vector of exogenous factors that influence centrality and can be constant or vary across countries) allows for a positive centrality even in the case of disconnected components of the graph.

In spite the fact that the Bonacich centrality allows for a more flexible formulation,<sup>26</sup> and the wide use of  $C_B$  in applications of network analysis in economics (Zenou, 2012) we do not calculate it, since it is a linear transformation of  $C_E$ .

We produce instead a weighted version of the out and in-eigenvector centrality measures.  $C_{E_{out}}^{NW}$  and  $C_{E_{in}}^{NW}$  are summarized in column (7) and (8) in table 2 and discussed in section 3.2.

### 3.2. Global trade through centrality indexes lenses

As we have seen in the previous paragraph, each index focuses on a particular aspect of the concept of centrality. In the empirical analysis, we need to make a choice between the different measures of centrality based on the characteristics of the phenomenon and the network we are looking at.

Showing a ranking of index values (calculated on the whole set of 178 countries), it is possible to identify which countries are more integrated in World Trade Network and the differences emerging from the use of different indicators. In Table 3 and 4 we report the ranking calculated from the same centrality indices included in table 1 and 2. For simplicity, we only report the ranking for a limited number of countries; in particular, we choose to focus on OECD, BRICS and “next eleven” (N-11) countries, identified as the most prominent countries in international trade. We report and comment the ranking related to 2007 indices value, the same reference year of the one considered in the figures included in section 2 and 3. Furthermore, our comment focuses on the top positions in the ranking. When interesting, we look at the weighted and unweighted version of indices to highlight differences in the position of a country in a network emerging when we explicitly consider the size of trade flows.

As highlighted in the previous paragraph, the Out-degree and In-degree measures (columns (1) and (2) in table 3) and the Out-degree percent and In-degree percent (columns (3) and (4) in table 3) respectively differ only in terms of the normalization factor (see equation 3 and 4), showing a correlation of 1 and an identical country ranking. Also the Out-closeness

<sup>26</sup> See the original paper by Bonacich (1987) for an interpretation of  $\alpha$  and  $\beta$  and Newman (2010) and Jackson (2010) for an overview. Kali and Reyes (2007) exploit the flexibility of the measure to include as exogenous  $\alpha$  factors the level of openness of the country or its GDP per capita. There is no generally valid procedure to follow and the selection of  $\alpha$  should be driven by the scope of the analysis.

and the In-closeness unweighted indices (columns (5) and (6) in table 3) show an identical ranking than the previous two indices. This is not a general result, it is due in the specific case of the aggregated World Trade Network in 2007 to the presence of a high number of strongly connected countries. In other words, if many countries are directly connected with all other  $N - 1$  countries in the network, the number of steps in the shortest path from a country and its possible partners is close to one (see equation 7): the Pearson correlation between  $C_{D_{out}}^N$  and  $C_{C_{out}}^N$  is 0.98, while the rank correlation is perfect. Only in the case of the eigenvector centrality the ranking in table 1 changes. Not dramatically in the top players ranking; much more in the case of countries in intermediate ranking positions.

Regarding weighted indicators, Out-strength and In-strength measures (columns (1) and (2) in table 4) and Out-strength and In-strength percent measures (columns (3) and (4) in table 4) also show a Pearson and a Spearman correlation of 1, respectively, differing only in terms of the normalizing factor (see equation 5 and 6). Closeness and Eigenvector centrality measures show instead a different ranking. The linear correlation between  $C_{S_{out}}^P$  and  $C_{C_{out}}^{NW}$  is 0.39, while the linear correlation between  $C_{S_{out}}^P$  and  $C_{E_{out}}^{NW}$  is 0.86.

Generally, when we look at unweighted centrality measures, the integration process in world trade is enhanced and the increasing importance of regional trade and of some emergent countries become evident. Instead, when we look at weighted centrality measures, a more traditional centre-periphery picture is confirmed.

Looking at tables 3 and 4 measure by measure, the in-degree index  $C_{D_{in}}^N$  shows that the big European countries are the most integrated in the World Trade Network. Germany, France, UK, Spain, Italy and the small open economy of Netherlands occupy the first place in the ranking, while China, Mexico and US the second one. However, when we look at the export side,  $C_{D_{out}}^N$  shows that not only China and the US rank first, but also several Asian countries, like India and Indonesia, display the highest index value: a clear evidence of the increased importance of South East Asia in international trade. It is interesting to notice that also smaller countries (Belgium, Switzerland and Denmark in the case of Europe) or emerging ones (Brazil and Turkey) rank second, showing an high number of direct link with the rest of the world. When we look at weighted version of these indexes, not surprisingly  $C_{S_{in}}^N$  shows that the US is the more central country in the World Trade Network. European countries, like Germany, France, UK, Spain, Italy and the Netherlands are placed in the top 10 positions, confirming their centrality in trade relationships, especially in terms of imports, due to their high degree of economic development. China ranks third.

As for weighted closeness,  $C_{C_{in}}^N$  highlights the importance of US centrality in the world trade. In the first four positions we can find Canada (2) and Mexico(4), two countries strictly tied to US through the NAFTA Agreement. China rank sixth, but it jumps to the first position when we look at the Out-closeness index  $C_{C_{out}}^N$ . Since eigenvector centrality

measures look at the country's centrality in terms of the centrality of its neighbors,  $C_{Ein}^N$  shows that Indonesia and India, together with China, rank first, reflecting the importance of the countries from which they import. However their ranking is much lower in the case of  $C_{Eout}^N$  (27, 14 and 6 respectively), reflecting their strong trade linkages with other Asian countries, whose neighbors are less central on a global scale.

As previously mentioned, when we look at the weighted eigenvector index, a more traditional picture emerges. US rank first in terms of  $C_{Eout}^{NW}$ , followed by Canada and Germany. Among developing countries, only China (4) is included in the leading group: again, a country whose trade relationship with the US (and other developed countries) is particularly strong. This last evidence is more clear when we look at In-eigenvector index  $C_{Ein}^{NW}$ , where China ranks first. Like the unweighted version of this index, and differently from other centrality measures, the importance of China's neighbors in terms of imports is higher than in terms of export flows.

**Table 3 – Centrality - unweighted measures**

id	i	iso3	country	(1) Out-degree	(2) In-degree	(3) Out-degree percent	(4) In-degree percent	(5) Out-closeness	(6) In-closeness	(7) Out-eigenvector	(8) In-eigenvector
10	36	AUS	Australia	3	6	3	6	3	6	9	19
11	40	AUT	Austria	1	6	1	6	1	6	1	21
16	50	BGD	Bangladesh	11	15	11	15	11	15	25	40
21	58	BLX	Belgium-Luxembourg	2	5	2	5	2	5	2	16
25	76	BRA	Brazil	2	7	2	7	2	7	6	24
30	124	CAN	Canada	3	3	3	3	3	3	4	11
32	152	CHL	Chile	21	38	21	38	21	38	45	80
33	156	CHN	China	1	2	1	2	1	2	1	6
42	203	CZE	Czech republic	7	3	7	3	7	3	15	9
45	208	DNK	Denmark	2	8	2	8	2	8	2	22
49	818	EGY	Egypt	15	24	15	24	15	24	32	54
51	233	EST	Estonia	23	46	23	46	23	46	47	100
53	246	FIN	Finland	3	14	3	14	3	14	5	34
55	251	FRA	France	1	1	1	1	1	1	1	3
43	276	DEU	Germany	1	1	1	1	1	1	1	4
65	300	GRC	Greece	9	13	9	13	9	13	20	31
72	348	HUN	Hungary	6	16	6	16	6	16	12	41
78	352	ISL	Iceland	32	41	32	41	32	41	58	85
74	699	IND	India	1	4	1	4	1	4	1	14
73	360	IDN	Indonesia	1	10	1	10	1	10	1	27
76	364	IRN	Iran	37	59	37	59	37	59	69	118
75	372	IRL	Ireland	6	5	6	5	6	5	13	17
79	376	ISR	Israel	10	17	10	17	10	17	29	47
80	381	ITA	Italy	1	1	1	1	1	1	1	1
83	392	JPN	Japan	2	3	2	3	2	3	3	13
104	484	MEX	Mexico	7	2	7	2	7	2	18	8
120	528	NLD	Netherlands	1	1	1	1	1	1	1	2
123	554	NZL	New Zealand	6	13	6	13	6	13	17	33
118	566	NGA	Nigeria	30	14	30	14	30	14	57	36
121	579	NOR	Norway	5	17	5	17	5	17	11	42
125	586	PAK	Pakistan	4	12	4	12	4	12	10	32
128	608	PHL	Philippines	9	25	9	25	9	25	23	58
130	616	POL	Poland	6	4	6	4	6	4	14	15
132	620	PRT	Portugal	8	17	8	17	8	17	19	43
136	643	RUS	Russian Federation	10	16	10	16	10	16	22	38
147	703	SVK	Slovak Republic	13	17	13	17	13	17	27	45
148	705	SVN	Slovenia	19	31	19	31	19	31	37	66
175	711	ZAF	South Africa	3	3	3	3	3	3	8	10
89	410	KOR	South Korea	2	5	2	5	2	5	7	18
50	724	ESP	Spain	2	1	2	1	2	1	2	5
149	752	SWE	Sweden	1	7	1	7	1	7	1	23
31	757	CHE	Switzerland	2	3	2	3	2	3	2	12
159	792	TUR	Turkey	2	10	2	10	2	10	2	26
165	842	USA	United States	1	2	1	2	1	2	1	7
58	826	GBR	United kingdom	1	1	1	1	1	1	1	4
170	704	VNM	Vietnam	14	31	14	31	14	31	30	63

To sum up, when we look at the unweighted centrality measures, the increased importance of some Asian countries is highlighted. A more traditional picture, however, emerges when

**Table 4 – Centrality - weighted measures**

id	i	iso3	country	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Out-strength	In-strength	Out-strength percent	In-strength percent	W-Out-closeness	W-In-closeness	W-Out-eigenvector	W-In-eigenvector
10	36	AUS	Australia	25	22	25	22	26	25	21	19
11	40	AUT	Austria	24	23	24	23	15	14	26	20
16	50	BGD	Bangladesh	71	69	71	69	71	72	62	75
21	58	BLX	Belgium-Luxembourg	11	9	11	9	13	12	13	10
25	76	BRA	Brazil	22	29	22	29	28	28	24	25
30	124	CAN	Canada	9	12	9	12	5	2	5	2
32	152	CHL	Chile	38	47	38	47	47	51	36	43
33	156	CHN	China	1	3	1	3	1	6	1	4
42	203	CZE	Czech Republic	31	28	31	28	19	21	32	30
45	208	DNK	Denmark	33	30	33	30	39	29	38	32
49	818	EGY	Egypt	63	49	63	49	80	57	66	50
51	233	EST	Estonia	76	74	76	74	97	90	92	88
53	246	FIN	Finland	34	37	34	37	50	35	42	36
55	251	FRA	France	5	5	5	5	9	8	9	7
43	276	DEU	Germany	2	2	2	2	2	3	4	3
65	300	GRC	Greece	61	39	61	39	78	42	69	40
72	348	HUN	Hungary	36	31	36	31	30	26	40	33
78	352	ISL	Iceland	98	95	98	95	108	124	95	90
74	699	IND	India	20	17	20	17	27	27	25	21
73	360	IDN	Indonesia	27	34	27	34	29	43	22	37
76	364	IRN	Iran	39	46	39	46	41	53	39	52
75	372	IRL	Ireland	26	33	26	33	25	23	23	29
79	376	ISR	Israel	48	45	48	45	35	47	44	38
80	381	ITA	Italy	6	7	6	7	12	9	11	11
83	392	JPN	Japan	4	6	4	6	4	7	3	8
104	484	MEX	Mexico	14	14	14	14	7	4	6	6
120	528	NLD	Netherlands	8	8	8	8	11	11	12	9
123	554	NZL	New Zealand	59	56	59	56	65	58	58	54
118	566	NGA	Nigeria	41	51	41	51	22	64	37	51
121	579	NOR	Norway	30	35	30	35	33	41	28	35
125	586	PAK	Pakistan	65	53	65	53	70	63	63	59
128	608	PHL	Philippines	40	43	40	43	49	52	35	41
130	616	POL	Poland	29	20	29	20	21	17	31	22
132	620	PRT	Portugal	49	38	49	38	52	31	54	39
136	643	RUS	Russian Federation	12	18	12	18	24	18	16	18
147	703	SVK	Slovak Republic	44	44	44	44	44	38	49	47
148	705	SVN	Slovenia	60	58	60	58	63	55	64	62
175	711	ZAF	South Africa	37	36	37	36	48	44	34	34
89	410	KOR	South Korea	10	13	10	13	8	16	7	13
50	724	ESP	Spain	16	10	16	10	23	13	18	14
149	752	SWE	Sweden	21	24	21	24	38	22	27	26
31	757	CHE	Switzerland	19	19	19	19	16	15	17	17
159	792	TUR	Turkey	32	21	32	21	40	30	43	23
165	842	USA	United States	3	1	3	1	3	1	2	1
58	826	GBR	United Kingdom	7	4	7	4	10	5	10	5
170	704	VNM	Vietnam	46	42	46	42	45	37	45	42

we take into account the size of bilateral trade flows. Looking at weighted indices, the result is a framework that emphasizes the importance of US and European countries as key players in international trade (degree and strength indices,) the centrality of US as a hub in the trade network (closeness, eigenvector,) the importance of China as a partner of the major advanced countries (eigenvector).

If we extend the analysis, including each of the 178 countries in the dataset, significant differences in the local and global ranking position emerges for countries less integrated in the World Trade Network. Figure 9 plots together the rankings obtained from local (out-degree and out-strength) and global (out-eigenvector) indices. The greater the distance from the diagonal, the greater the difference in the relative positioning of the country expressed by the two indices. In particular, countries above (below) the diagonal are less (more) central when their position is measured by a local with respect to a global measure of centrality.

Panel (a) in figure 9 shows that, when unweighted centrality measure are calculated,

rankings obtained from global and local indices are quite similar. As we mentioned before, this is due to the presence of a high number of strongly connected countries in the World Trade Network.<sup>27</sup> However, in panel (b) of figure 9 (weighted indicators), a significant dispersion is associated with the central part of the ranking positions, while it is lower at the two extremes. It follows, that the major difference between *local centrality measures* and *global centrality measures* concerns those countries showing an intermediate position in the country ranking. Countries above the diagonal increase their ranking position when centrality includes the “effect of others,” i.e. the structural component influencing their position in the network.

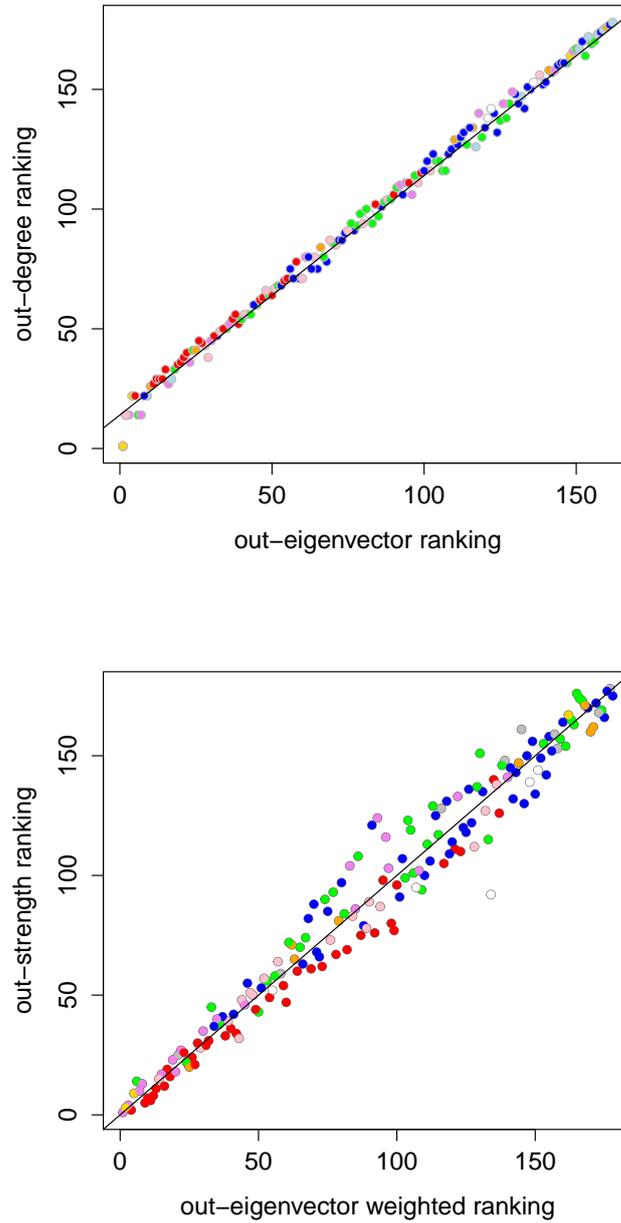
Like in the figures of previous paragraphs, colors depict geographical regions. European countries (red) are mainly located below the diagonal (more in details, those European countries are Eastern European peripheral countries such as the Baltic States, Cyprus, Ukraine, Bosnia, and Turkey,) while above the diagonal it is possible to find a group of Latin American and African countries (green and blue). The “below-the-diagonal” European countries are therefore less central from a structural point of view, even if they are closely related to EU countries. More generally, if a country is closely related to a very integrated area (like the EU) but this relationship represents almost all of its overall bilateral trade, the transition from local to global measures of centrality shows a significant reduction of its importance in the World Trade Network ranking. Similarly, African and especially Latin American countries appear to be more structurally connected at the global level.

Among the countries above the diagonal, it is also possible to identify a number of countries whose peripheral location within the network depends on geographical (islands such as Aruba, Madagascar, Jamaica, Cuba, Haiti, New Caledonia) or political (North Korea) factors. Their ties with central partners increase their own global centrality, which is much lower than that determined by the weight of their bilateral trade and/or the number of their partners.

---

<sup>27</sup> The position of the “diagonal” line has been rescaled to take into account that 13 countries are sharing the first position in the out-degree ranking.

**Figure 9 –**  
Local and global centrality, 2007. Respectively unweighted and weighted network



**Note:** Since we are confronting ranked measure, in both panels more central countries are closer to the origin. Different colors correspond to different geographical regions. In panel (a) of the figure the origin of the 45° line has been rescaled so to take into account that 13 countries were ranking first in the Out-degree centrality measure. Trade data come from BACI-CEPII dataset.

## 4. Sectoral Trade Networks: Bananas, Cement, Movies, Oil, Footwear and Engines

One of the great advantages of the BACI-CEPII database is the possibility of entering into bilateral trade flows in great detail. The data used since now is an aggregation of more than 5000 sectoral data, which is now worth looking at using the sociograms and the data selection procedure used for figure 3.

Since the scope of this section is just to scratch the surface of the network analysis of sectoral data,<sup>28</sup> we have selected some sectors according to some specific product characteristics or to their relevance in the world trade. The sectors we have selected are: Bananas (Hs6-1992: 80300, Bananas, including plantains, fresh or dried), Cement (Hs6-1992: 252329, Portland cement), Movies (Hs6-1992: 370610, Cinematograph film, exposed and developed, width > 35m), Oil (Hs6-1992: 270900, Petroleum oils, oils from bituminous minerals, crude), Footwear (Hs6-1992: 640399, Footwear, sole rubber, plastics uppers of leather) and Engines (Hs6-1992: 870323, Automobiles, spark ignition engine of 1500-3000 cc). We follow the same compression scheme that we followed in figure 2 and figure 3, selecting only the major two import flows for every country, and considering the binary (unweighted) version of the resulting trade matrix. The visualization of each country is proportional to its out-degrees.

### 4.1. Bananas

International trade in bananas has been a topic of both academic research (Josling and Taylor, 2003) and policy debate (Anania, 2009). Characterized by preferential schemes, strong international conflict and an important role of fair-trade independent agencies, according to FAO, bananas are one of the most important foods for both consumption and trade. More than 81 million tonnes of bananas were produced in 2007, 20% of them are exported. They are the fourth most important food staple in the world and the fifth most-traded agricultural commodity (after cereals, sugar, coffee and cocoa), generating billions of dollars of earnings.

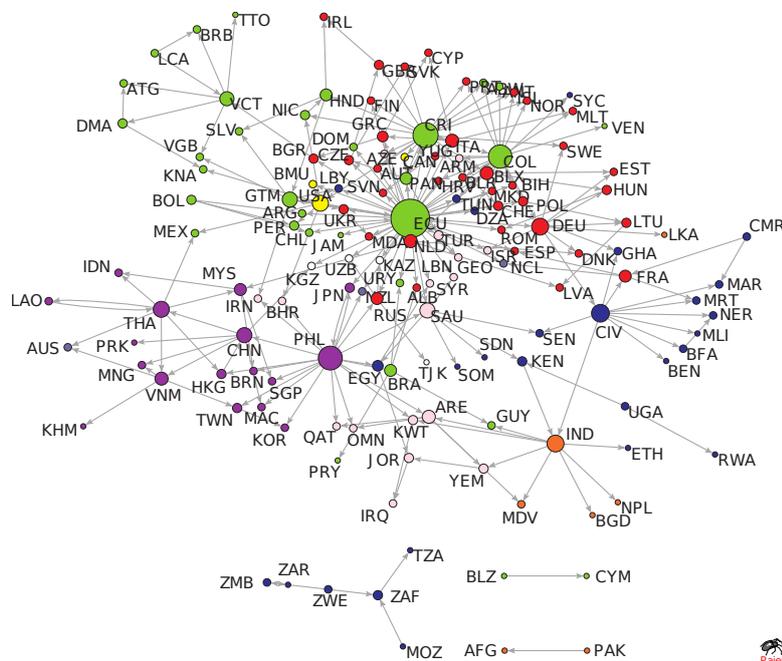
In figure 10 for every country in the network we recorded only the first two importing partners in terms of import flows in 2007. The size of the circle associated to each country is proportional to the country's out-degree.

International trade in bananas looks geographically segmented, with the major role played by Asia, on one side, and Latin America, on the other. The Asian regional trade is dom-

---

<sup>28</sup> For a more systematic sectoral network analysis see De Benedictis and Tajoli (2010) and Barigozzi, Fagiolo and Garlaschelli (2010), where the structure of the sectoral network and its topological properties are related to the complexity of the good traded internationally. Some network analyses of specific sectors are Amighini and Gorgoni (2013) and Blázquez and González-Díaz (2012) on the automotive sector, Akerman and Larsson (2010) on arms trade, Beretta, Maggioni, Uberti and Beretta (2013) on miscellaneous goods.

**Figure 10 –**  
World Trade Network in Bananas (major two importing partners, 2007)



**Note:** Only the import flows from the first and second trade partner are taken into consideration for each country. 30 disconnected countries are not included. Country labels are the iso3 country codes (see the Appendix for a complete list of the countries included in the analysis). The size of the circle associated to each country is proportional to out-degrees. Different colors correspond to different geographical regions. Trade data come from BACI-CEPII dataset. The network is drawn using Pajek.

inated by the Philippines (PHL), with China (CHN), Thailand (THA), Malaysia (MYS) and Vietnam (VNM) as followers, while the Latin American sectoral trade is dominated by Ecuador (ECU), with Costa Rica (CRI) and Colombia (COL) in a supportive position. North American and especially European countries (the main sectoral importers together with China, Japan and Russia) are clearly integrated with the Latin American trade sub-network, with the partial exception of France (FRA) that acts as a bridge with the African sub-network. Africa, with the leading role played by Côte d'Ivoire (CIV), and India (the main producer of bananas, by far) operate in a decentralized position.

Finally, two cases are noteworthy. Brazil (BRA) is neatly separated from the rest of Latin American countries and orbits around the Asian trade system by the attractive role played by Egypt (EGY). On the other extreme of figure 10, Saint Vincent and the Grenadines (VCT) plays the role of a hub in the small islands bananas trade system.

The sectoral trade system is also characterized by 30 isolated nodes (not shown in figure

10) and by three isolated small sub-systems in Africa (with South Africa (ZAF) as a hub), South America and South Asia.

Apart from the country-by-country relations, the sector is characterized by an oligopolistic structure where six or seven very large firms dominate the market. This could also further explain, together with the trade preferences guaranteed by the European Union to the members of the Africa, Caribbean and Pacific (ACP) trade partnership, some of the features in figure 10.

#### 4.2. Cement

After water, the second most consumed material in the world is concrete. Economic growth and concrete are cemented together: modern buildings and public infrastructures, such as roads, bridges, dams, schools, sewers and hospitals are made of concrete. In spite of that, the cement and concrete industry has been widely studied because of the homogeneity and the very limited spatial mobility of the products in question. (Syverson, 2004) (Syverson, 2006) focuses on ready-mixed concrete, the fluid compound made of cement, sand, gravel, water, and some chemical additives. The product “has a low value-to-weight ratio and is highly perishable - it absolutely must be discharged from the truck before it hardens.” Proximity to customers is therefore unavoidable and “... foreign trade in ready-mixed concrete is essentially nonexistent.” In fact, in the BACI-CEPII Hs (or Site) classification there is no code corresponding to ready-mixed concrete. Looking for an as hardly tradable product than concrete, we now examine the World Trade Network of the main ingredient of it: Portland cement.

Often referred to as OPC (Ordinary Portland Cement,) Portland cement is the most common type of cement in general use around the world. It is a fine powder produced by grinding cement clinker (between 90 and 95%) and gypsum and up to 5% of minor constituents. The cement clinker usually originates from limestone, and that’s make it a “local good”: it is heavy, cheap, and limestone is largely available on local markets.<sup>29</sup>

Cement is produced in 156 countries, and the global production in 2007 is around 2700 million tonnes. The supply of cement is however not homogeneously distributed across countries. China accounts for half of the world’s production and India is a distant second, with a production of 183 million tonnes. The top ten producing countries account for 70% of total cement production.

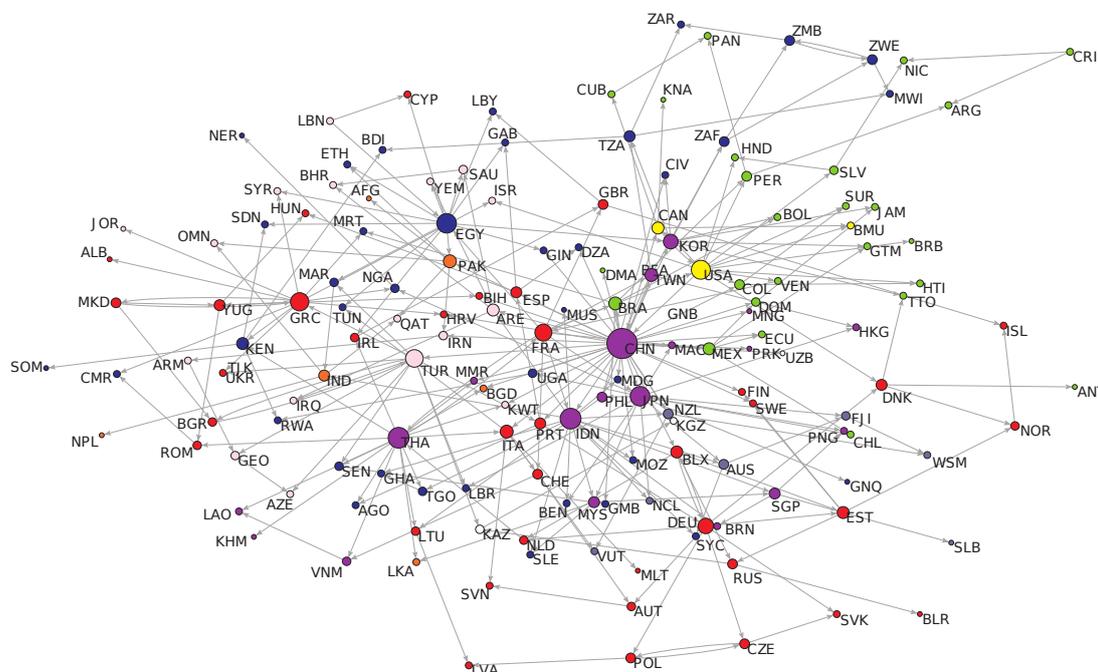
Portland cement is indeed an hardly tradable product: 95% of world production is consumed domestically.<sup>30</sup> That is reflected in the picture emerging from figure 11, which is

---

<sup>29</sup> The market for OPC is however far from perfect competition. Domestic markets are highly concentrated, incumbent firms often form cartels (Levenstein and Suslow, 2006; Röller and Steen, 2006) and high capital costs and the access to limestone reserves act as an entry barrier for potential entrants.

<sup>30</sup> Germany is the major exporter of Portland cement, followed closely by Belgium and China; the US are the major importer, followed by The Netherlands and France.

**Figure 11 –**  
World Trade Network in Portland Cement (major two importing partners, 2007)



**Note:** Only the import flows from the first and second trade partner are taken into consideration for each country. Country labels are the iso3 country codes (see the Appendix for a complete list of the countries included in the analysis). The size of the circle associated to each country is proportional to out-degrees. Different colors correspond to different geographical regions. Trade data come from BACI-CEPII dataset. The network is drawn using Pajek.

quite different from that in figure 10. Though China and other Asian countries, such as Japan (JPN), Indonesia (IND) and Thailand (THA), stand out from the crowd, there is no major evidence of any central players. Trade is dominated by distance. Bilateral links are especially among adjacent countries; continental links are neatly separated in the case of Latin American countries, less so in the case of African or European countries, even if some patterns are evident: a “mediterranean” sub-network is clustered at the north-east of figure 11, incorporating India and Pakistan (linked to Egypt and Morocco, in turn); and a Francophone Sub-Saharan Africa clustered at the south-east of the figure.

In general, it is as if continental trade links were superimposed one on top of the others, with no countries acting as a predominant attractors. This is the effect of the little relevance of international trade in the sector, together with the fundamental role of intra-firm trade by construction services multinationals, and non-tariff barriers (USITC, 2011).

### 4.3. Movies

Trade in cultural goods (books, recorded music, painting, movies and the like)<sup>31</sup> is rapidly growing, at a rate which is more than five times the average rate of growth of total global trade (UNESCO, 2005). Trade is remarkably concentrated, being in the hands of a handful of countries, and in some cases, such the one of the US, cultural products is an export industry of comparable size with respect to traditional manufacturing industries. Moreover, trade in cultural goods is also studied as a way of quantifying countries' cultural proximity (Disdier *et al.*, 2009).<sup>32</sup> In figure 12 we focus on movies only. The centrality of the US is more than evident. It is the global hub of the trade network in movies. The existence of regional clusters is also evident. Italy (ITA), France (FRA), the UK (GBR) and also Russia have a predominant position in the European trade system, exporting to francophone and anglophone African countries, and to the Middle Est. Mexico (MEX) and Argentina (ARG) play the same role in Latin America, as India (IND) in Asia with Japan (JPN), and Thailand (THA), Korea (KOR) and Hong Kong (HKG) (not China) acting has bridges between the East and the West. Also the United Arab Emirates (ARE) acts as a regional hub. Language and cultural models are influencing the cluster formation, but, in any case, the trade system is not continentally segmented and only a minority of 32 countries (not included in figure 12) looks disconnected from the giant component of the network.<sup>33</sup>

### 4.4. Crude Oil

In the crude oil trade network Russia (RUS) plays a central role in the European trade system, with Kazakhstan (KAZ) acting as a secondary hub in Central Asia. The country acts as a bridge between Europe an China.

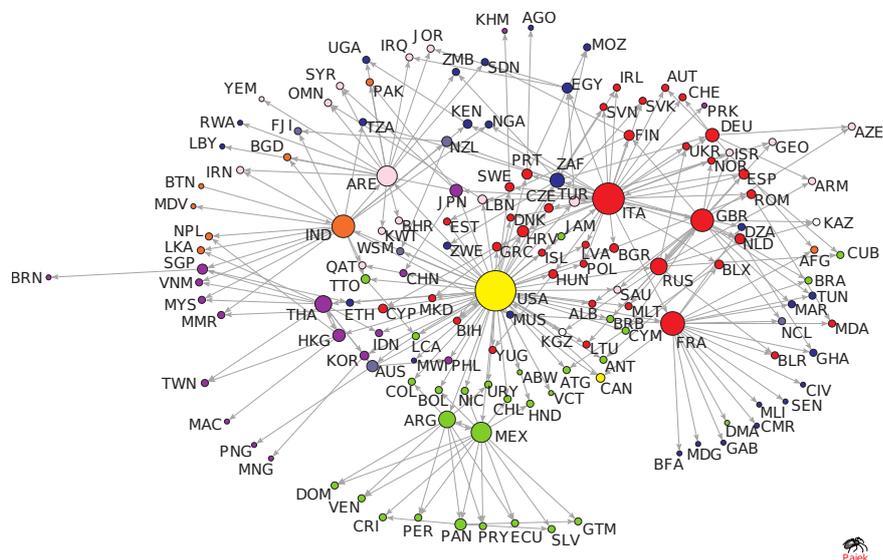
---

<sup>31</sup> The definition of cultural goods is time-varying and lacks consensus. Recently the UNESCO (2005) provided a new definition of cultural goods to take account the revolution in ITC, modern art and social habits. Cultural goods and services are “... tangibles and intangibles conveying cultural content that might take either the form of a good or a service, [and also] the goods and services which are required to produce and disseminate such content [...] as well as ancillary services even if they are only partly cultural in their content.” Disdier, Tai, Fontagne' and Mayer (2009) uses this definition (and the BACI-CEPII database,) focusing on goods.

<sup>32</sup> A further relevant question is the one of the consequences of globalization in determining a convergence in social habits or imposing a dominant “way of life”. Olivier, Thoenig and Verdier (2008) shows that the reduction of national varieties associated with the trade-induced “selection effect,” and the resulting cultural convergence, can counterbalance the welfare effect of goods market integration. Francois and van Ypersele (2002) shows that “national exception” protectionist measures are welfare enhancing if consumers love variety and firms care about the scale of production.

<sup>33</sup> Marvasti and Canterbury (2005) investigates US exports in motion pictures at the beginning of the '90s, finding for a sample of 33 countries a positive impact of language, education and religion on export flows. Hanson and Xiang (2011) uses data on trade in movies to quantify global exports fixed costs and their prevalence over bilateral export fixed costs.

**Figure 12 –**  
World Trade Network in Movies (major two importing partners, 2007)



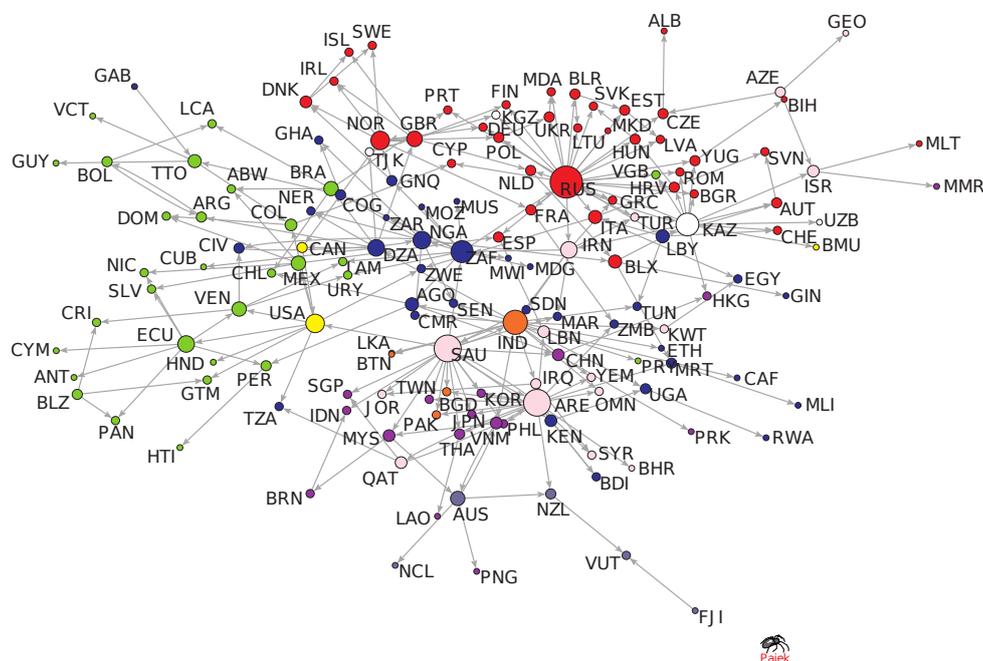
**Note:** Only the import flows from the first and second trade partner are taken into consideration for each country. 32 disconnected countries are not included. Country labels are the iso3 country codes (see the Appendix for a complete list of the countries included in the analysis). The size of the circle associated to each country is proportional to out-degrees. Different colors correspond to different geographical regions. Trade data come from BACI-CEPII dataset. The network is drawn using Pajek.

In recent years, China has been investing heavily in Kazakhstan's oil and in the country's infrastructure, including pipelines, highways and rail roads, and refineries. China's purchases of Kazakhstan's oil, gas and minerals has booming, and Hong Kong is the gateway to China.

Latin American countries appears as composed of two separate cohesive sub-groups. On one side, Venezuela (VEN), Mexico (MEX), Ecuador (ECU) coordinate the northern cluster and Brazil (BRA), Argentina (ARG) and Colombia (COL) coordinate the southern one, both groups of leading countries keep the intra-continent trade connected and they also link the continent to the rest of the world. Saudi Arabia (SAU) and the United Arab Emirates (ARE) are dominant in the Middle East and are projected towards Asia, being paced between China (CHN) and Japan (JAP). It is worth noting the regionally decentralized position of Iran (IRN), gravitating towards Europe.

India (IND) and the US (USA) also emerge as central played. Finally, African countries, such as Nigeria (NGA) and South Africa (ZAF), are placed in between Europe, Asia and South America, with Angola (AGO) and Algeria (DZA) playing a reinforcing position. The Oceanic regional system is the most peripheral one.

**Figure 13 –**  
World Trade Network in Oil (major two importing partners, 2007)



**Note:** Only the import flows from the first and second trade partner are taken into consideration for each country. 27 disconnected countries are not included. Country labels are the iso3 country codes (see the Appendix for a complete list of the countries included in the analysis). The size of the circle associated to each country is proportional to out-degrees. Different colors correspond to different geographical regions. Trade data come from BACI-CEPII dataset. The network is drawn using Pajek.

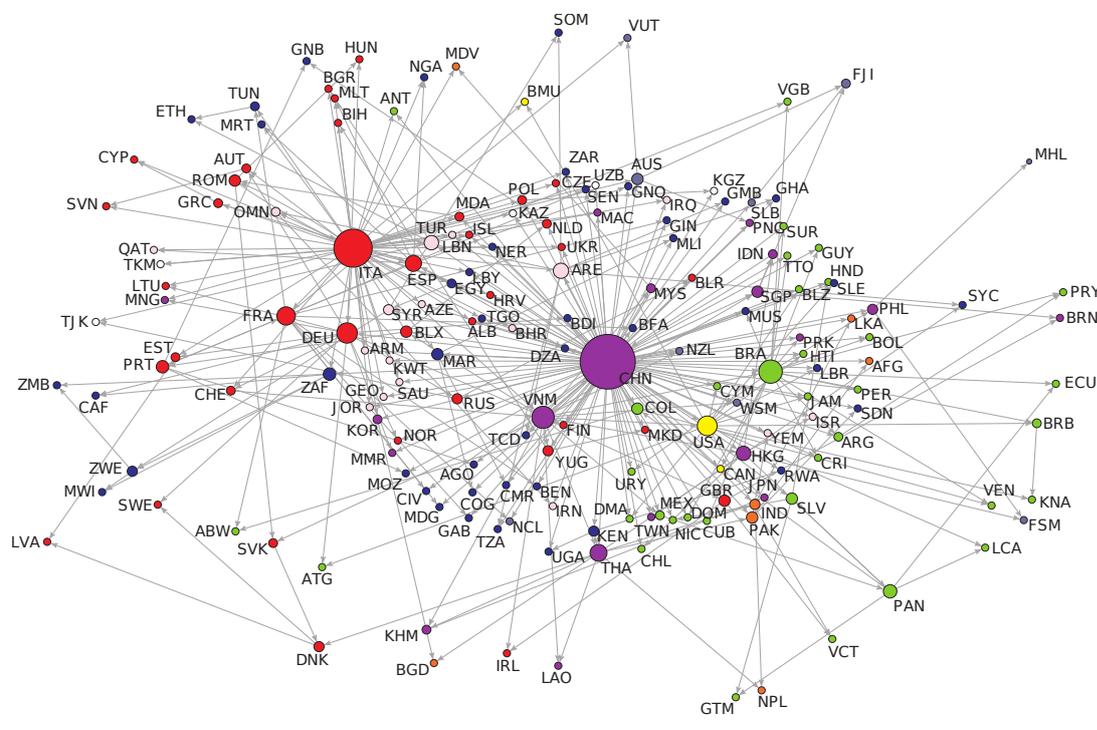
#### 4.5. Footwear

The footwear sector is one of the most important sector in international trade. In 2007, 12 billions of pairs of shoes were traded internationally and the trend is still growing. In spite of the theory of the international product cycle, the sector has not been fully transferred to new industrialized countries: 9 out of top 15 exporters are European countries, still. On the other hand, China is the largest exporter, by far (with some 35% of total world exports).

Figure 14 shows how the footwear network is polarized around China (CHN) and Italy (ITA), acting as the loci of the sectoral trade system. Among the other European countries, Germany (DEU), France (FRA) and Spain (ESP) look prominent, followed by Portugal (PRT), Romania (ROM) and Denmark (DNK). The UK (GBR) has a totally decentralized position with respect to the rest of Europe.

Thailand (THA) and especially Vietnam (VNM) are the satellites of China (CHN) in the

**Figure 14 –**  
World Trade Network in Footwear (major two importing partners, 2007)



**Note:** Only the import flows from the first and second trade partner are taken into consideration for each country. 1 disconnected country is not included. Country labels are the iso3 country codes (see the Appendix for a complete list of the countries included in the analysis). The size of the circle associated to each country is proportional to out-degrees. Different colors correspond to different geographical regions. Trade data come from BACI-CEPII dataset. The network is drawn using Pajek.

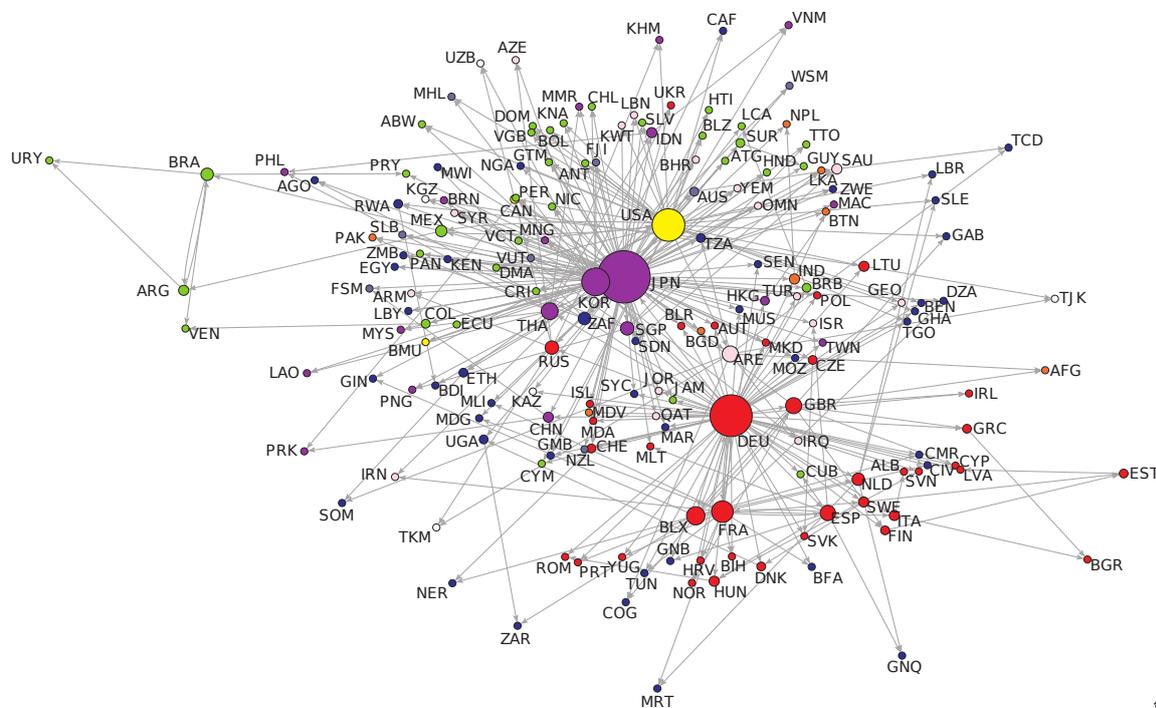
Asiatic trade system, with Hong Kong in an intermediate position between the US (USA), India (IND) and Pakistan (PAK) and Latin American countries, such as Brazil (BRA), El Salvador (SLV) and Colombia (COL), with Panama (PAN) at the extreme periphery.

A bird-eye view of the sectoral network show a much more dense cluster in the case of China's international relations, while for Italy it looks more sparse.

#### 4.6. Engines

A similar topological structure can be observed also in the Automobile Engine sector, another sector of major importance in international trade. Germany (DEU) acts as an European focal node, while Japan (JPN) is the Asiatic one. The picture is made a little bit more complex by the role of subsequent major players. The US (USA) acts as a third pole, more close Japan than to Germany. The Asian system is also influence by the

**Figure 15 –**  
World Trade Network in Engines (major two importing partners, 2007)



**Note:** Only the import flows from the first and second trade partner are taken into consideration for each country. Country labels are the iso3 country codes (see the Appendix for a complete list of the countries included in the analysis). The size of the circle associated to each country is proportional to out-degrees. Different colors correspond to different geographical regions. Trade data come from BACI-CEPII dataset. The network is drawn using Pajek.

role of South Korea (KOR), Thailand (THA) and Singapore (SGP). Notwithstanding its major role of a producer, the position of China (CHN) in the Engine trade networks looks marginal, close to the Russian Federation (RUS) and orbiting around the European cluster.

In the European trade system, we can notice the role of the UK (GBR), France (FRA) and Spain (SPA) and of some new European member countries, such as the Czech Republic (CZE), Hungary (HUN) and Lithuania (LTU), exploiting their convenient geographical location and lower labor costs to boost international fragmentation of production from old European member countries.

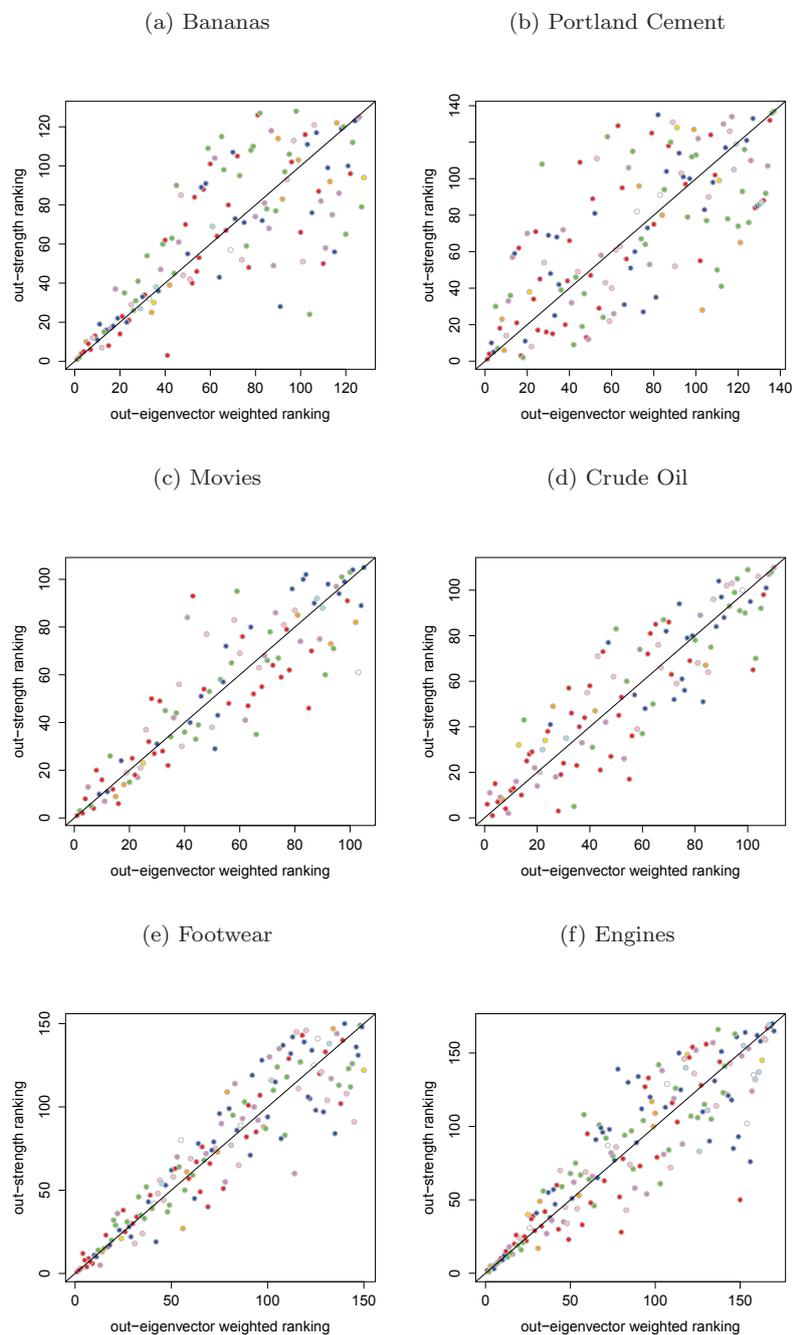
Finally, while Latin American countries are in general closer to the Japan and US orbit, it is worth noting the peculiarity of the sub-network, constituted by Brazil (BRA), Argentina (ARG) and Uruguay (URY) which will be disconnected from the rest of the global sectoral network if it were not for the link with Mexico (MEX) and, through it, with the US.

#### 4.7. Centrality in sectoral data

It is of course possible to calculate centrality measures starting from sectoral data. Using a visual comparison similar to the one of figure 9, we can compare *local* and *global* centrality measures. For each of the six sectors discussed in the previous session, figure 16 plots together the rankings of out-degree and out-eigenvector measures for each of the 178 countries of our dataset. An overall outlook at figure 16 shows a quite important level of dispersion common to all sectors, although different for each one. Like in the case of figure 9, it means that the relative position of a country in a network changes if we take into consideration local rather than global centrality measures. In some cases, this dispersion is less evident in the lower part of the diagonal (see for instance the bananas, footwear and engines sectors). It follows that the main gap between local and global centrality measures occurs for those countries less integrated in the world trade, i.e. holding a low position in the centrality ranking of the sectoral network, both in terms of Out-strength and Out-eigenvector indices. Conversely, countries playing an important role in sectoral trade - both as main players or neighbors of them - and holding a key position in the centrality ranking, do not display any difference between local and global measures. The dispersion is associated to geographical regions as well. Differently from figure 9, countries belonging to the same continent are not clustered together above or below the diagonal but hold a scattered position, and this is true in all sectors selected.

A specific analysis for single sector provides additional insights. The Portland cement sector (figure 16(b)) - which has the most fragmented network structure, as showed in figure 11 - presents the highest level of dispersion all along the diagonal. It means that countries that are key players at local level become less important when the "effect of others" (i.e. the structural component) is taken into consideration, and vice versa. On the contrary, the Footwear sector (figure 16(e)) characterized by a dense and duopoly structure, as seen in figure 14 - shows the lowest level of dispersion compared to the other sectors. In this case, a relative wider dispersion below the line denotes that countries are less central when their position is measured by a global centrality index with respect to a local one.

**Figure 16** –  
Strength and Eigenvector centralities at the sectoral level, 2007.



**Note:** Only the import flows from the first and second trade partner are taken into consideration for each country. Country labels are the iso3 country codes (see the Appendix for a complete list of the countries included in the analysis). The size of the circle associated to each country is proportional to out-degrees. Different colors correspond to different geographical regions. Trade data come from BACI-CEPII dataset. The network is drawn using Pajek.

## 5. Conclusions

In this paper we explore the World Trade Network through the lens of Network Analysis introducing the reader to some of the techniques used to visualize, calculate and synthetically represent network trade data. In particular, we show different visualizations of the network and describe its topological properties, producing and discussing some of the commonly used Network's statistics, and presenting some specific topics. To do this, we make use of the BACI dataset; this latter, constantly updated by the CEPII and largely available, offers the richness of the UN Comtrade database in terms of number of countries, large period of time, disaggregated product level, but with the advantage of significantly reducing the number of missing values. We show that Network Analysis is a useful tool to describe bilateral trade relations among countries, characterized by high dimensionality and strong heterogeneity. This latter is highlighted and measured in terms of the position of each country in the trade network according to its relative centrality.

Centrality measures can be *local*, taking into considerations only the direct links of a country (i.e. its neighborhood) or *global* (when we consider the effect of third countries). These latter, in particular, are of some interest to show on the one hand the strong changes in trade relations occurred in the past decades, on the other hand the complexity of the structure of World Trade Network. However, if the majority of countries are strictly connected to each other, their position in terms of centrality as drawn from local and global indicators tends to coincide. In dense networks the global position of a country can be inferred by its local one.

It follows that the global indicators are particularly suitable for the analysis of countries with an intermediate degree of integration. Their relative position within the World Trade Network can vary greatly depending on the neighbors centrality, the type of trade flow considered (global or sectoral flows), as well as the type of bilateral relationship, in its unweighted (i.e. binary) and weighted (e.g. trade flows) version.

In particular, looking at unweighted centrality measures, the increased importance of some Asian countries has been highlighted. A more traditional picture, however, has emerged in the case of weighted indices. The result is a framework that emphasizes the importance of US and European countries as a player in international trade (degree and strength indices), the centrality of US as a hub in the World Trade Network (closeness, eigenvector), the importance of China as a partner of the major advanced countries (eigenvector).

Looking at sectoral data, the network visualization of some selective products puts emphasis - once more - on the extreme heterogeneity of data, showing markets mainly characterized by an oligopolistic structure and remarkably polarized networks.

We just scratched the surface of possible analyses that can be done using the BACI-CEPII dataset in a network framework. For instance, the interested reader can easily replicate the same analysis that we have done in this paper, comparing years before and after the

financial crisis and the great trade collapse of the end of 2008. The 1995-2010 centrality measures for the 178 countries in the dataset, which is available for downloading at the [CEPII databases page](#) allow an immediate comparison. Moreover, the **Stata**, **R** and **Pajek** scripts included in the Appendix 7 make it possible to apply the same techniques used in this paper to different dataset, so as to encourage the empirical trade analyst to explore the possibilities offered by network analysis. To conclude, the paper suggests that a network approach can be fruitfully applied in several contexts of international trade analysis. From the seminal paper by Casella and Rauch (2002), the interest of international trade economists to the role played by networks of international relations as evolved, bringing together theory and empirical analysis (Goyal and Joshi, 2006; Chaney, 2011). The focus on the structural dimension of trade relations and on the interdependence among countries, that is the starting point of Network Analysis, can nicely complement other empirical analyses (such as the gravity model). This very approach can be used to address issues where systemic effects can be relevant, such as the transmission mechanisms of international shocks, the role of the extensive and the intensive margins in trade dynamics, the relationship between multilateralism and regionalism, and the impact of new emerging countries in the World Trade Network. Furthermore, network indices can be applied in an innovative manner to the study of country economic performance, providing complementary information to the one of the most commonly used measures in empirical international trade analysis.

## 6. Appendix: Data and indicators

Table 5 – Iso3 Country codes, Country names and Geographical Regions

iso3	Country	Continent/Geographical region	Color	iso3	Country	Continent/Geographical region	Color
ABW	Aruba	Latin America	Green	KWT	Kuwait	Middle East	Pink
AFG	Afghanistan	South Asia	Orange	LAO	Laos	East Asia	Purple
AGO	Angola	Africa	Blue	LBN	Lebanon	Middle East	Pink
ALB	Albania	Europe	Red	LBR	Liberia	Africa	Blue
ANT	Netherlands Antilles	Latin America	Green	LBY	Libya	Africa	Blue
ARE	United Arab Emirates	Middle East	Pink	LCA	Saint Lucia	Latin America	Green
ARG	Argentina	Latin America	Green	LKA	Sri Lanka	South Asia	Orange
ARM	Armenia	Middle East	Pink	LTU	Lithuania	Europe	Red
ATG	Antigua and Barbuda	Latin America	Green	LVA	Latvia	Europe	Red
AUS	Australia	Australia	Navy	MAC	Macau	East Asia	Purple
AUT	Austria	Europe	Red	MAR	Morocco	Africa	Blue
AZE	Azerbaijan	Middle East	Pink	MDA	Moldova	Europe	Red
BDI	Burundi	Africa	Blue	MDG	Madagascar	Africa	Blue
BEN	Benin	Africa	Blue	MDV	Maldives	South Asia	Orange
BFA	Burkina Faso	Africa	Blue	MEX	Mexico	Latin America	Green
BGD	Bangladesh	South Asia	Orange	MHL	Marshall Islands	Australia	Navy
BGR	Bulgaria	Europe	Red	MKD	Macedonia	Europe	Red
BHR	Bahrain	Middle East	Pink	MLI	Mali	Africa	Blue
BIH	Bosnia and Herzegovina	Europe	Red	MLT	Malta	Europe	Red
BLR	Belarus	Europe	Red	MMR	Myanmar	East Asia	Purple
BLX	Belgium-Luxembourg	Europe	Red	MNG	Mongolia	East Asia	Purple
BLZ	Belize	Latin America	Green	MOZ	Mozambique	Africa	Blue
BMU	Bermuda	North America	Yellow	MRT	Mauritania	Africa	Blue
BOL	Bolivia	Latin America	Green	MUS	Mauritius	Africa	Blue
BRA	Brazil	Latin America	Green	MWI	Malawi	Africa	Blue
BRB	Barbados	Latin America	Green	MYS	Malaysia	East Asia	Purple
BRN	Brunei Darussalam	East Asia	Purple	NCL	New Caledonia	Australia	Navy
BTN	Bhutan	South Asia	Orange	NER	Niger	Africa	Blue
CAF	Central African Republic	Africa	Blue	NGA	Nigeria	Africa	Blue
CAN	Canada	North America	Yellow	NIC	Nicaragua	Latin America	Green
CHE	Switzerland	Europe	Red	NLD	Netherlands	Europe	Red
CHL	Chile	Latin America	Green	NOR	Norway	Europe	Red
CHN	China	East Asia	Purple	NPL	Nepal	South Asia	Orange
CIV	Cote d'Ivoire	Africa	Blue	NZL	New Zealand	Australia	Navy
CMR	Cameroon	Africa	Blue	OMN	Oman	Middle East	Pink
COG	Congo	Africa	Blue	PAK	Pakistan	South Asia	Orange
COL	Colombia	Latin America	Green	PAN	Panama	Latin America	Green
CRI	Costa Rica	Latin America	Green	PER	Peru	Latin America	Green
CUB	Cuba	Latin America	Green	PHL	Philippines	East Asia	Purple
CYM	Cayman Islands	Latin America	Green	PNG	Papua New Guinea	East Asia	Purple
CYP	Cyprus	Europe	Red	POL	Poland	Europe	Red
CZE	Czech Republic	Europe	Red	PRK	North Korea	East Asia	Purple
DEU	Germany	Europe	Red	PRT	Portugal	Europe	Red
DMA	Dominica	Latin America	Green	PRY	Paraguay	Latin America	Green
DNK	Denmark	Europe	Red	QAT	Qatar	Middle East	Pink
DOM	Dominican Republic	Latin America	Green	ROM	Romania	Europe	Red
DZA	Algeria	Africa	Blue	RUS	Russian Federation	Europe	Red
ECU	Ecuador	Latin America	Green	RWA	Rwanda	Africa	Blue
EGY	Egypt	Africa	Blue	SAU	Saudi Arabia	Middle East	Pink
ESP	Spain	Europe	Red	SDN	Sudan	Africa	Blue
EST	Estonia	Europe	Red	SEN	Senegal	Africa	Blue
ETH	Ethiopia	Africa	Blue	SGP	Singapore	East Asia	Purple
FIN	Finland	Europe	Red	SLB	Solomon islands	Australia	Navy
FJI	Fiji	Australia	Navy	SLE	Sierra Leone	Africa	Blue
FRA	France	Europe	Red	SLV	El Salvador	Latin America	Green
FSM	Micronesia	Australia	Navy	SOM	Somalia	Africa	Blue
GAB	Gabon	Africa	Blue	SUR	Suriname	Latin America	Green
GBR	United Kingdom	Europe	Red	SVK	Slovak Republic	Europe	Red
GEO	Georgia	Middle East	Pink	SVN	Slovenia	Europe	Red
GHA	Ghana	Africa	Blue	SWE	Sweden	Europe	Red
GIN	Guinea	Africa	Blue	SYC	Seychelles	Africa	Blue
GMB	Gambia	Africa	Blue	SYR	Syria	Middle East	Pink
GNB	Guinea-Bissau	Africa	Blue	TCD	Chad	Africa	Blue
GNQ	Equatorial Guinea	Africa	Blue	TGO	Togo	Africa	Blue
GRC	Greece	Europe	Red	THA	Thailand	East Asia	Purple
GTM	Guatemala	Latin America	Green	TJK	Tajikistan	Central Asia	White
GUY	Guyana	Latin America	Green	TKM	Turkmenistan	Central Asia	White
HKG	Hong Kong	East Asia	Purple	TTO	Trinidad and Tobago	Latin America	Green
HND	Honduras	Latin America	Green	TUN	Tunisia	Africa	Blue
HRV	Croatia	Europe	Red	TUR	Turkey	Middle East	Pink
HTI	Haiti	Latin America	Green	TWN	Taiwan	East Asia	Purple
HUN	Hungary	Europe	Red	TZA	Tanzania	Africa	Blue
IDN	Indonesia	East Asia	Purple	UGA	Uganda	Africa	Blue
IND	India	South Asia	Orange	UKR	Ukraine	Europe	Red
IRL	Ireland	Europe	Red	URY	Uruguay	Latin America	Green
IRN	Iran	Middle East	Pink	USA	United States	North America	Yellow
IRQ	Iraq	Middle East	Pink	UZB	Uzbekistan	Central Asia	White
ISL	Iceland	Europe	Red	VCT	Saint Vincent and the Grenadines	Latin America	Green
ISR	Israel	Middle East	Pink	VEN	Venezuela	Latin America	Green
ITA	Italy	Europe	Red	VGB	Virgin Islands	Latin America	Green
JAM	Jamaica	Latin America	Green	VNM	Vietnam	East Asia	Purple
JOR	Jordan	Middle East	Pink	VUT	Vanuatu	Australia	Navy
JPN	Japan	East Asia	Purple	WSM	Samoa	Australia	Navy
KAZ	Kazakhstan	Central Asia	White	YEM	Yemen	Middle East	Pink
KEN	Kenya	Africa	Blue	YUG	Serbia and Montenegro	Europe	Red
KGZ	Kyrgyzstan	Central Asia	White	ZAF	South Africa	Africa	Blue
KHM	Cambodia	East Asia	Purple	ZAR	Congo Dem. Rep.	Africa	Blue
KNA	Saint Kitts and Nevis	Latin America	Green	ZMB	Zambia	Africa	Blue
KOR	South Korea	East Asia	Purple	ZWE	Zimbabwe	Africa	Blue

Note: The column “color” includes the color corresponding to each Geographic Region in the paper’s figures. See figure 2 as an example.

## 7. Appendix: R, Stata and Pajek routines

### 7.1. Appendix: Calculating networks cumulative distributions in R

The following R-script allows to import in R the data in Pajek format. The dataset `b2007.net` including 2007 aggregated bilateral trade flows from the BACI-CEPII database is in Pajek net format; importing it in R requires the `igraph` library. The `read.graph` transforms the Pajek format in an `igraph` graph object.

```

R code
library(igraph)
w2007_full <- read.graph("~/data/b2007.net", format="pajek")
summary(b2007)
```

The `summary` command generates the following outcome:

```
IGRAPH D-W- 178 22002 --
attr: id (v/c), x (v/n), y (v/n), z (v/n), weight (e/n)
```

IGRAPH indicates that this is an `igraph` graph. The subsequent four bits denote the graph: the first ‘D’ is for directed graphs (‘U’ for undirected). The second, blank in our case, is ‘N’ if the graph has the ‘name’ vertex attribute set. The third is ‘W’ for weighted graphs, if the ‘weight’ edge attribute is set. The fourth is ‘B’ for bipartite graphs. The subsequent two numbers, indicate the number of vertices (i.e. 178) and the number of arcs in the directed graph (i.e. 22002), and after a double dash, eventually, the name of the graph. Now we can use the data to generate the in-degree and out-degree cumulative distributions as in figure 5.

```

R code

dd.in <- degree.distribution(w2007_full, mode="in", cumulative=TRUE)
dd.out <- degree.distribution(w2007_full, mode="out", cumulative=TRUE)

plot(dd.in, xlab="in-degree and out-degree", ylab="cumulative frequency", col=2, type="l", main="",
xlim=c(1,210), ylim=c(0.01,1.2)) \

lines(dd.in, col="light grey", type="p", cex=0.8, pch=4)
lines(dd.in, col="2", type="l")
points(177,dd.in[177], col="2", type="p", cex=0.8, pch=1)

text(196, 0.03, "US / China / Italy", srt = 0, font = 3, col = 2, cex = 0.7) #Add text; "srt" rotates
to 0 degrees, font = 3 == italic; degrees: 177 \

points(92,dd.in[92], col="2", type="p", cex=0.8, pch=1)
text(102, dd.in[92], "Congo", srt = 0, font = 3, col = 2, cex = 0.7) # degrees: 92
text(134, dd.in[145], "Tunisia", srt = 0, font = 3, col = 2, cex = 0.7) # degrees: 145
points(145,dd.in[145], col="2", type="p", cex=0.8, pch=1)
text(26, 1.04, "Micronesia", srt = 0, font = 3, col = 2, cex = 0.7) # degree: 24
points(24,dd.in[24], col="2", type="p", cex=0.8, pch=1)
text(119, dd.in[129], "Uganda", srt = 0, font = 3, col = 2, cex = 0.7) # degree: 129
points(129,dd.in[129], col="2", type="p", cex=0.8, pch=1)

lines(dd.out, col=4, type="l")
lines(dd.out, col="light grey", type="p", cex=0.8, pch=4)
lines(dd.out, col="4", type="l")

par(fig=c(0.07, 0.5, 0.07, 0.5), cex=0.6, new = T)
plot(dd.in, log="xy", col=2, type="l", xlab="", ylab="")
lines(dd.in, log="xy", col="light gray", type="p", cex=0.6, pch=4, xlab="", ylab="", axes=FALSE)
lines(dd.in, log="xy", col="2", type="l")
lines(dd.out, log="xy", col=4, type="l", xlab="", ylab="")

par(fig=c(0.0, 1, 0.0, 1), cex=1, new = T)
```

## 7.2. Appendix: A brief note on Stata routines for network centrality measures

The following commands allows to reproduce in Stata 12 the centrality measures included in the paper. Some of them are straightforward (degree and strength,) and therefore are not included. The others require a little bit of manipulation, that we guide through the following do file.

```

----- Stata code -----
*** Closeness ***
/* Since it is based on an adjacency matrix, the Stata SGL package consider the value of the link as distance by default;
an higher trade flow among two countries is interpreted as higher distance,
that is why we need to normalize and take the inverse of the trade flow when computing adjacency based centrality measures */

summ w
g w_average = w/r(mean)
g inv_w = 1/w_average

* directed closeness defined over Clos-OUT
netsis i j, measure(distance) name(A, replace) weight(inv_w) directed
netsummarize (rows(A)-1):rowsum(A), generate(clossw_out) statistic(rowsum)

* directed closeness defined over Columns: Clos-IN
netsis i j, measure(distance) name(Aaa, replace) weight(inv_w) directed
netsummarize (cols(Aaa)-1):colsum(Aaa), generate(clossw_in) statistic(colsum)

/* Note that in order to compute In/Out centrality measures for directed (weighted or unweighted) networks you need
to calculate the indexes over Columns or Rows, here an example for closeness, but it works the same also
for degree and strength. */

*** Eigenvector centrality ***
/* When compute the eigenvector centrality for directed (weighted or unweighted) SGL package reports by default the
right leading eigenvector, measuring the relative importance of the nodes that point to the considered vertex.
In order to get the left eigenvector just invert the vertex order ij -> ji. */

*IN
netsis i j, measure(eigenvector) name(E, replace) directed weight(w)
netsummarize E, generate(eigenw_in) statistic(rowsum)

*OUT
netsis j i, measure(eigenvector) name(E, replace) directed weight(w)
netsummarize E, generate(eigenw_out) statistic(rowsum)

```

## 7.3. Appendix: Pajek routines for network visualization

The following commands allows to reproduce in Pajek (3.07 version) the visualizations showed in Figs. 2, 3, 6, 7, and 8. Visualizations on your computer screen could be sometimes slightly different from the figures in the paper, mainly because of different personal settings. Be sure the general patterns match only.

```

----- Pajek code -----

In order to replicate Figure 2:

"File -> Pajek Project File -> Read -> Figure_2.paj"
The project file load in the associated work space a Network a Partition and a Vector.
Note that: Partitions store discrete characteristics of the vertices (in this case continents clusters);
Vectors stores continuous attributed of each node (in this case Indegree).
The Network files contains an array of the links among nodes (and their labels),
moreover in this version the file contains also the geographical coordinates of each node.
"Network -> Create Vector -> Centrality ->## "
In order to produce other vectors with alternative centrality measures:

```

```

"Draw -> Network"
Will draw the graph for the network loaded into Pajek memory -- first slot of the network section.
Note: coordinates need to be scaled to a range [0,1].
"Draw -> Network + First Partition"
Now the resulting graph associates a different color to each category of the partition file
(in such a case continents). \\

"Draw -> Network+ First Partition+ First Vector"
Visualize Network with colors for each continents and node size proportional to vector attribute
(in-degree in this case). \\

In order to replicate Figure 3:
"File -> Pajek Project File -> Read -> Figure_3.paj"
Drawing commands are the same of Figure_2, but in this case the xy coordinates of each node are
those produced by one of the optimization algorithms build in Pajek (namely the Kamada-Kawai force algorithm).
Note that each time you run the Kamada-Kawai, as well as other optimization procedures, the coordinates
of the single node may be different, the overall relation (distance) between nodes does not change,
only the visualization does; if you want to replicate exactly the graph you see you need to save
the network leaving the graph window open. In such a way Pajek will save the network with
the xy coordinates used to produce the graph.

In order to replicate Figure 6:
Graph 6A:
"File -> Pajek Project File -> Read -> Figure_6A.paj"
The project file loads in the associated work space a Network (the exports flows among 178 countries in 2007)
and a Partition (the continents clusters). \\

"Network -> Create Vector->Centrality->Degree->Output" will create a vector of out-degree flows
In order to narrow the size of network to 25\% of world flows (i.e. larger flows):
"Network -> Create new network -> Transform -> Remove -> Lines with value -> Lower than->#"
"Network -> Create partition -> Valued core -> First threshold and step -> All"
Type the # in the dialog box asking for the threshold and type the highest flows value in a second dialog box.
"Operations-> Vector+Partition-> Extract subvector-> 1-*" will extract the vector of countries whose
flows are above the threshold. \\

"Operations->Network + Partition-> Extract Subnetwork -> 1-*" will create the new network of countries
whose flows are above the threshold.\\

"Partitions ->Extract sub-partition ->1-*" to maintain different colors of continents
(make sure that the original continents clusters and the created partitions below and above
the threshold are selected as the first and second partition respectively in the Partitions menu).
Now load Vector named "OutdegreeMainPlayers" storing the out-degrees referred to the main 17 players.

"Draw-> Network+ First Partition + First Vector" to visualize Network with colors for each continents
and node size proportional to vector attribute (out-degree in this case).
Note: nodes size needs to be scaled.

Graph 6B:
"File -> Pajek Project File -> Read -> Figure_6B.paj"
The project file loads the same Network and Partition of Fig.6A and a Vector which stores GDP values.
In order to narrow the size of network to 50\% of world flows:
"Network -> Create new network -> Transform -> Remove -> Lines with value -> Lower than->##"
Repeat the same commands as Fig.6A:
"Network -> Create partition -> Valued core -> First threshold and step -> All"
Type the ## in the dialog box asking for the threshold and type the highest flows value in a second dialog box.
"Operations-> Vector+Partition-> Extract subvector-> 1-*"
"Operations->Network + Partition-> Extract Subnetwork -> 1-*"
"Partitions ->Extract sub-partition ->1-*" to maintain different colors of continents (make sure that the
original continents clusters and the created partitions below and above the threshold are selected as the
first and second partition respectively in the Partitions menu).

"Draw-> Network+ First Partition + First Vector"
Note: nodes size needs to be scaled.

Graph 6C:
"File -> Pajek Project File -> Read -> Figure_6C.paj"
The project file loads the same Network and Partition of Fig.6A and a Vector which stores per
capita GDP values. In order to narrow the size of network to 75\% of world flows, follow the same
command of Fig.6B selecting the corresponding threshold ###

Graph 6D:
"File -> Pajek Project File -> Read -> Figure_6D.paj"
The project file loads the same Network and Partition of Fig.6A and a Vector which stores openness indexes
(i.e. export to GDP). In order to narrow the size of network to 90\% of world flows though maintaining all
countries (i.e. included nodes with links value below the threshold):

"Network -> Create new network -> Transform -> Remove -> Lines with value -> Lower than->####"
"Draw-> Network+ First Partition + First Vector" to visualize the whole Network with colors for
each continents and node size proportional to vector attribute (openness in this case).
Note: nodes size needs to be scaled.

```

In order to replicate Figure 7:

```
"File -> Pajek Project File -> Read -> Figure_7.paj"
```

The project file loads a Network (the exports flows among 178 countries in 2007), a Partition (the continents clusters) and a Vector (the per capita GDP attribute).

```
"Operations->Network + Partition->Extract SubNetwork"
```

and, because Europe is the third class in the partition, select class 3 in the dialog box that appears.

Select again the original continent partition in the Partition drop-down menu then:

```
"Operations->Vector + Partition->Extract Subvector"
```

and select again 3 in the dialog box.

Before drawing the network be sure that the new subnetwork is selected as the first in each drop-down menu of the Main screen.\\

```
"Draw ->Network + First Partition + First Vector"
```

If necessary, energize it to obtain a drawing similar to Figure 7 choosing the Kamada-Kawai->Free command from the Layout->Energy menu (this command produces regularly spaced results, especially for connected networks that are not very large).

In order to replicate Figure 8:

```
"File -> Pajek Project File -> Read -> Figure_8.paj"
```

The project file loads a Network (the exports flows among 178 countries in 2007) and a Partition (the continents clusters).\\

```
"Operations->Network + Partition->Shrink Network"
```

and select 1 in the dialog box asking for the minimum number of connections between clusters and accept the default value (zero) in a second dialog box.

```
"Draw->Network + First Partition"
```

To get a picture more similar to Figure 8 some manipulations are needed. Pajek chooses the label of the first vertex of a class that is shrunk and adds a pound sign (#) to obtain a label for the shrunken class. For example, Albania happens to be the first European country in the network, so the vertex that represents this continent carries the label "#ALB" in the shrunken network. We added the names of continents to Figure 8 manually by editing the shrunken partition with the File->Partition->Edit command.

Loops are not visible in the Draw screen: they appear once exported the drawing as a picture.

## References

- Akerman A. and Larsson A. (2010) The global arms trade network 1950-2007, *Stockholm University Working Paper*, 2.
- Amighini A. and Gorgoni S. (2013) The international reorganization of auto production, *The World Economy*, forthcoming.
- Anania G. (2009) How would a wto agreement on bananas affect exporting and importing countries?, *ICTSD Programme on Agricultural Trade and Sustainable Development, International Centre for Trade and Sustainable Development*, 21.
- Anderson J.E. and vanWincoop E. (2003) Gravity with gravitas, *American Economic Review*, 93, 170–192.
- Barigozzi M., Fagiolo G. and Garlaschelli D. (2010) The Multi-Network of International Trade: A Commodity-Specific Analysis, *Physical Review E*, 81, 046104.
- Beretta S., Maggioni M.A., Uberti E.T. and Beretta S. (2013) Books, butter and guns: networks and distances in international trade, *mimeo*.
- Blázquez L. and González-Díaz B. (2012) International automotive production networks: How the web comes together, *mimeo*.
- Bonacich P. (1972) Factoring and weighting approaches to status scores and clique identification, *The Journal of Mathematical Sociology*, 2, 1, 113–120.
- Bonacich P. (1987) Power and centrality: A family of measures, *The American Journal of Sociology*, 92, 5, 1170–1182.
- Bonacich P. and Lloyd P. (2001) Eigenvector-like measures of centrality for asymmetric relations, *Social Networks*, 23, 191–201.
- Borgatti S.P. (2005) Centrality and network flow, *Social networks*, 27, 1, 55–71.
- Bowen H., Hollander A. and Viaene J. (2012) *Applied International Trade*, Palgrave Macmillan.
- Butts C.T. (2008a) network: A Package for Managing Relational Data in R, *Journal of Statistical Software*, 24, 2, 1–36.
- Butts C.T. (2008b) Social Network Analysis with sna, *Journal of Statistical Software*, 24, 6, 1–51.
- Cartwright D. and Harary F. (1977) A graph theoretic approach to the investigation of system-environment relationships, *Journal of Mathematical Sociology*, 5, 87–111.
- Casella A. and Rauch J.E. (2002) Anonymous Market and Group Ties in International Trade, *Journal of International Economics*, 58, 1, 19–47.
- Chaney T. (2011) The Network Structure of International Trade, *NBER wp*, 16753.
- Clauset A., Shalizi C. and Newman M. (2009) Power-Law Distributions in Empirical Data, *SIAM Review*, 51, 4, 661–703.
- Combe D., Langeron C., Egyed-Zsigmond E. and Géry M. (2010) A comparative study of

- social network analysis tools, *International Workshop on Web Intelligence and Virtual Enterprises*.
- Csardi G. and Nepusz T. (2006) The igraph software package for complex network research, *International Journal of Complex Systems*, 1695.
- Davis D.R. and Weinstein D.E. (2001) What Role for Empirics in International Trade?, in: *Bertil Ohlin: A Centennial Celebration, 1899-1999*, Findlay, Jonung and Lundahl, eds., MIT Press, Cambridge.
- De Benedictis L. and Tajoli L. (2010) Comparing Sectoral International Trade Networks, *Aussenwirtschaft*, 65, 167–189.
- De Benedictis L. and Tajoli L. (2011) The World Trade Network, *The World Economy*, 34, 8, 1417–1454.
- de Nooy W., Mrvar A. and Batagelj V. (2011) *Exploratory Social Network Analysis with Pajek*, Structural Analysis in the Social Sciences, Cambridge University Press.
- Dijkstra E.W. (1959) A note on two problems in connexion with graphs, *Numerische Mathematik*, 1, 269–271.
- Disdier A.C., Tai S.H.T., Fontagne L. and Mayer T. (2009) Bilateral trade of cultural goods, *Review of World Economics*, 1, 269–271.
- Fagiolo G., Reyes J. and Schiavo S. (2008) On the Topological Properties of the World Trade Web: A Weighted Network Analysis, *Physica A: Statistical Mechanics and its Applications*, 387, 15, 3868–3873.
- Feenstra R.C. and Taylor A.M. (2008) *International Trade*, Palgrave Macmillan Limited.
- Felbermayr G. and Koheler W. (2005) Exploring the Extensive and Intensive Margins of World Trade, *Review of World Economics*, 142, 4, 642–674.
- Fisman R. and Wei S.J. (2004) Tax Rates and Tax Evasion: Evidence from “Missing Imports” in China, *Journal of Political Economy*, 112, 2, 471–500.
- Fisman R. and Wei S.J. (2009) The Smuggling of Art, and the Art of Smuggling: Uncovering the Illicit Trade in Cultural Property and Antiques, *American Economic Journal: Applied Economics*, 1, 3, 82–96.
- Francois P. and van Ypersele T. (2002) On the protection of cultural goods, *Journal of International Economics*, 56, 2, 359–369.
- Fruchterman T. and Reingold E. (1991) Graph Drawing by Force-directed Placement, *Software - Practice and Experience*, 21, 11, 1129–1164.
- Gabaix X. (2009) Power Laws in Economics and Finance, *Annual Review of Economics*, 1, 255–93.
- Gabaix X. (2011) The Granular Origins of Aggregate Fluctuations, *Econometrica*, 79, 733–772.
- Garlaschelli D. and Loffredo M.I. (2005) Structure and Evolution of the World Trade Network, *Physica A: Statistical Mechanics and its Applications*, 355, 138–144.

- Gaulier G. and Zignago S. (2010) BACI: International Trade Database at the Product-level - The 1994-2007 Version, *CEPII Working Paper*, 23, 1–28.
- Goodreau S.M., Handcock M.S., Hunter D.R., Butts C.T. and Morris M. (2008) A statnet Tutorial, *Journal of Statistical Software*, 24, 9, 170–192.
- Goyal S. (2009) *Connections: An Introduction to the Economics of Networks*, Princeton University Press.
- Goyal S. and Joshi S. (2006) Bilateralism and Free Trade, *International Economic Review*, 47, 3, 749–778.
- Hanson G. and Xiang C. (2011) Trade Barriers and Trade Flows with Product Heterogeneity: An Application to U.S. Motion Picture Exports, *Journal of International Economics*, 83, 2, 14–26.
- Helpman E. and Krugman P. (1985) *Market structure and foreign trade: increasing returns, imperfect competition and the international economy*, Mit Press.
- Helpman E., Melitz M. and Rubinstein Y. (2008) Estimating Trade Flows: Trading Partners and Trading Volumes, *Quarterly Journal of Economics*, 123, 2, 441–487.
- Jackson M. (2010) *Social and Economic Networks*, Princeton University Press, Princeton University Press.
- Javorcik B.S. and Narciso G. (2008) Differentiated Products and Evasion of Import Tariffs, *Journal of International Economics*, 76, 208–222.
- Josling T.E. and Taylor T.G. (2003) *Banana Wars. The Anatomy of a Trade Dispute*, Wallingford: CABI Publishing.
- Kali R. and Reyes J. (2007) The Architecture of Globalization: A Network Approach to International Economic Integration, *Journal of International Business Studies*, 38, 4, 595–620.
- Kamada T. and Kawai S. (1989) An algorithm for drawing general undirected graphs, *Information processing letters*, 31, 12, 7–15.
- Katz L. (1953) A new status index derived from sociometric analysis, *Psychometrika*, 18, 1, 39–43.
- Levenstein M.C. and Suslow V.Y. (2006) What determines cartel success?, *Journal of Economic Literature*, 44, 1, 43–95.
- Marvasti A. and Canterbury E.R. (2005) Cultural and other barriers to motion pictures trade, *Economic Inquiry*, 43, 1, 39–54.
- Miura H. (2012) Stata graph library for network analysis, *The Stata Journal*, 12, 1, 94–129.
- Moreno J. (1934) *Who shall survive?*, Beacon Press, New York.
- Newman M. (2010) *Networks: An Introduction*, OUP Oxford.
- Olivier J., Thoenig M. and Verdier T. (2008) Globalization and the dynamics of cultural identity, *Journal of International Economics*, 76, 2, 356–370.

- Opsahl T. (2009) *Structure and Evolution of Weighted Networks*, University of London (Queen Mary College), London, UK.
- R Development Core Team (2007) *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.
- Riccaboni M. and Schiavo S. (2010) Structure and growth of weighted networks, *New Journal of Physics*, 12.
- Röller L.H. and Steen F. (2006) On the workings of a cartel: Evidence from the norwegian cement industry, *The American economic review*, 96, 1, 321–338.
- Rotunno L. and Vézina P.L. (2012) Chinese networks and tariff evasion, *The World Economy*, 35, 12, 1772–1794.
- Scott J. (2000) *Social Network Analysis: A Handbook*, SAGE Publications.
- Syverson C. (2004) Market structure and productivity: A concrete example, *Journal of Political Economy*, 112, 6, 1181–1222.
- Syverson C. (2006) Markets: Ready-mixed concrete, *Journal of Economic Perspectives*.
- UNESCO (2005) *International flows of selected cultural goods and services, 1994-2003*, Montreal, UNESCO Institute for Statistics.
- USITC (2011) Gray portland cement and cement clinker from japan, *U.S. International Trade Commission*, Investigation No. 731-TA-461 (Third Review), Publication 4281.
- Vega-Redondo F. (2007) *Complex Social Networks*, Econometric Society Monographs, Cambridge University Press.
- Wasserman S. and Faust K. (1994) *Social Network Analysis: Methods and Applications*, Structural Analysis in the Social Sciences, Cambridge University Press.
- Zenou Y. (2012) Networks in Economics, *Cepr discussion papers*, 9021.