Random Walks on the World Input-Output Network

Carlo Piccardi[1], Massimo Riccaboni[2], Lucia Tajoli[3], Zhen Zhu[4]

1 Politecnico di Milano, Milan, Italy
2 IMT School for Advanced Studies Lucca, Lucca, Italy, and KU Leuven, Leuven, Belgium
3 Politecnico di Milano, Milan, Italy, Corresponding author: lucia.tajoli@polimi.it
4 IMT School for Advanced Studies Lucca, Lucca, Italy

Abstract

Modern production is increasingly fragmented across countries. To disentangle the world production system at sector level, we use the World Input-Output Database to construct the World Input-Output Network (WION) where the nodes are the individual sectors in different countries and the edges are the transactions between them. In order to explore the features and dynamics of the WION, in this paper we detect the communities in the WION and evaluate their significance using a random walk Markov chain approach. Our results contribute to the recent stream of literature analysing the role of global value chains in economic integration across countries, by showing global value chains as endogenously emerging communities in the world production system, and discussing how different perspectives produce different results in terms of the pattern of integration.

Keywords: random walks; Markov chains; community detection; input-output analysis; world input-output network.

1 Introduction

The global financial crisis of 2007-08 has brought into sharp focus the fact that our economy is intricately connected across countries and sectors. In public debate, the notion of "too big to fail" was soon complemented by the network concepts such as "too central to fail" or "too connected to fail." And it has been proposed that the regulation of financial institutions should be based on their network positions in the financial system rather than their sheer size [6, 3]. Moreover, with a broader production system where financial sectors and non-financial ones are interconnected by input-output relationships, the network perspective has also been used to investigate the origin of macroeconomic fluctuations [1, 4] and to simulate the impact of local shocks on the whole economy [10].

The application of network analysis to economics goes well beyond crisis studies. In its simplest form, a network is merely a collection of actors (or nodes) and the relationships (or edges) between them [43]. The economic systems that fit this description are ubiquitous. Indeed, network analysis has made novel contributions to such diverse fields as trade and global value chains [59, 18, 12, 23, 22, 56, 5, 67, 66, 68, 60, 11], technology and innovation [53, 19, 7, 46, 41, 58], and industrial organization [33, 55]. A common method of network analysis used in the above literature is to calculate some node-level network measures such as centralities [27, 43, 16] and relate them to other economic variables. Another central topic of network analysis is community detection, i.e., finding possible partitions of a network into communities (or clusters) with relatively strong internal but weak external connections [20, 21]. Community detection has provided valuable insights in understanding various complex phenomena in different realms, from the Internet to biological ecosystems, to economics. In particular, in economics, community detection is important to highlight the structural characteristics of higher-order interconnections between economic agents (countries, sectors, firms and consumers) which are key to understand the pattern of economic relations and the propagation of economic effects, as highlighted by [1]. For example, community analysis has been applied to shed light on the pattern of international economic integration [64, 24, 51, 8], to detect corporate connections [49], or to study correlations between stocks [45, 50].

Our paper detects the communities in the World Input-Output Network (WION) using a random walk Markov chain approach. The WION is constructed from the World Input-Output Database, i.e., the Global Multi-Regional Input-Output (GMRIO) tables, covering 35 sectors for each of the 40 countries (27 EU countries and 13 major countries in other regions) in the years from 1995 to 2011 [63]. Therefore, for each year, the nodes of the WION are the individual sectors in each country, and the edges, directed and weighted, are the input-output relationships between them, which is the value of intermediate goods purchased and sold by each sector in each country, representing the connections between production structures.

The use of input-output tables to better understand trade patterns and countries linkages has a long tradition in the economic literature, starting from the early studies by Leontief [37]. Input-output tables have been used extensively to assess the potential specialization of countries in trade and the general equilibrium impact of trade policies (see for example [25]), and more recently to evaluate the environmental impact of countries' production and trade [65]. In the last years, the use of inter-country input-output tables was revived by a new stream of literature [29, 34, 63], aimed at improving the measurement of countries' participation to global value chains and to highlight production linkages between countries. But to the best of our knowledge, very few studies applied techniques of community detection to GMRIO tables [5].

For the study of the communities in the WION, random walk based approaches are the most natural choice. Indeed, there is a close relationship between input-output systems and Markov chains. It is well known that an open input-output system can be modeled as an absorbing Markov chain, with the productive sectors as the transient states and the value-added contribution or final demand as the absorbing states [31, 26, 62, 15, 30]. Standard community detection methods, such as max-modularity [44, 20, 5], could obviously be used, but the above theoretical link would be lost, whereas the Markov chain approach is the most appropriate method of conducting community detection if one wants to bridge modern network science to input-output systems.

Among the several Markov chain based community detection strategies (e.g., [52, 57, 14]), we restrict our attention to the *Lumped Markov Chain* approach [48] because it explicitly provides an assessment of the quality of each community. That is, it returns each community along with its *persistence probability*, a measure of cohesiveness that can be compared with a pre-defined cutoff value. The results obtained can be intuitively interpreted, allowing also to assess the specific significance of the emerging communities. Moreover, it allows the evaluation of the cohesiveness of any subnetworks exogenously defined. This is especially useful in the context of the WION because we can examine the evolution of the cohesiveness of a subnetwork of interest (say, a country or a sector) over time.

The rest of the paper is organized as follows. Section 2 presents in detail the methodology employed in the analysis. First, the linkage between an open input-output system and an absorbing Markov chain is reviewed. Then, we convert the open input-output system to a regular Markov chain by either removing the external value-added contribution and final demand, or by endogenizing them as an extra sector. Finally, we review the community detection algorithm that we will exploit for our analysis. In Section 3 we construct the WION from the GMRIO tables and conduct the community detection. Results are reported and thoroughly discussed. Section 4 provides further discussion and concludes the paper.

2 Methods

2.1 Open Input-Output Systems as Absorbing Markov Chains

A classical input-output system combines the internal input-output relationships between sectors and the external information of value-added contribution and final demand. It is also called an *open* input-output system since the system is open to exogenous shocks from supply or demand side (i.e., changes in the external information). In the literature of input-output economics, the supply side analysis is based on the Ghosh model while the demand side analysis is based on the Leontief model [40]. The Ghosh model concerns how sectors pass on the changes in external production costs (e.g., increase in wages) proportionally to all the sectors in the country and to final demand. Hence, the key assumption behind the Ghosh model is that the market share structure of outputs for each sector is stable. On the other hand, the Leontief model concerns how sectors satisfy the changes in external final demand by requiring fixed proportions of inputs from all the sectors in the country and from primary value-added. Hence, the key assumption behind the Leontief model is that the technical requirement structure of inputs for each sector is stable. Briefly, below we show the equivalence between the Leontief model and an absorbing Markov chain. The Ghosh model can be analyzed in a similar fashion.

An n-sector economic system can be compactly modeled by the $n \times n$ *transaction matrix* $\mathbf{Z}$, together with the *value-added vector* $\mathbf{v}$, the *final demand vector* $\mathbf{f}$ and the *total output vector*, $\mathbf{x}$, expressing the value of total sector production. The entry Z_{ij} of the transaction matrix $\mathbf{Z}$ denotes a transaction between sector i and sector j, and can be interpreted as the amount of material flow (intermediate inputs usually measured in monetary value) supplied from sector i to sector j or, equivalently, as the amount of monetary flow paid from sector j to sector i. The value-added vector $\mathbf{v}^{\mathrm{T}} = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$ includes the contribution of primary factors required by production, such as skilled and unskilled labor, physical production capital, etc. The final demand vector $\mathbf{f}^{\mathrm{T}} = \begin{bmatrix} f_1 & \cdots & f_n \end{bmatrix}$ is the output not absorbed by intermediate sector use but rather by household consumption, government expenditure, etc. Finally, the total output vector $\mathbf{x}^{\mathrm{T}} = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}$ is, for each entry x_i, the total production of sector i, which meets the balance $x_i = \sum_j Z_{ij} + f_i = \sum_i Z_{ij} + v_i$. It is standard to assume the non-degeneracy condition $x_i > 0$ for all i, namely each sector has indeed non-zero production.

The *technical coefficients matrix* is defined as $\mathbf{A} = \mathbf{Z}\hat{\mathbf{x}}^{-1}$, where we use the "hat" symbol over a vector to denote the diagonal matrix with the elements of the vector on its diagonal: its entry A_{ij} represents the share of input from sector i out of sector j's total output. It is easy to show that $\mathbf{Ax} + \mathbf{f} = \mathbf{x}$, that is, for each sector, the total output equals its provision of inputs to other sectors, including itself, plus the final demand. Solving for $\mathbf{x}$ we get $\mathbf{x} = (\mathbf{I} - \mathbf{A})^{-1}\mathbf{f}$, where $\mathbf{I}$ is the identity matrix, and the so-called *Leontief inverse* is defined as

$$\mathbf{L} = (\mathbf{I} - \mathbf{A})^{-1}, \tag{1}$$

whose entry L_{ij} measures the change of output in sector i due to a one-unit change in final demand for sector j.

Finally, note that usually the input-output system described above assumes the economic system to be given by single country. But the same notation can also effectively represent a GMRIO system, a group of countries having domestic as well as international economic transactions. The only difference is that the latter further specifies the country for each element of the input-output matrix, i.e., i actually refers to a sector/country pair. We refer the interested reader to [63] for more details of the GMRIO system we use in this paper.

A Markov chain [31] on the other hand is a stochastic process where transitions from one state to another obey the Markov property, i.e., the probability distribution of the next state solely depends on the current state. The states of a Markov chain can be partitioned into transient and ergodic sets. Once a process leaves a transient set, it will never come back. And once a process enters an ergodic set, it never leaves. If a state is the only element in an ergodic set, the state is called absorbing. An *absorbing Markov chain* is such that all its ergodic states are absorbing. A canonical form of the transition matrix of an absorbing Markov chain, with s transient states and r absorbing states, is

$$\mathbf{P} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{R} & \mathbf{Q} \end{bmatrix}, \tag{2}$$

where $\mathbf{I}$ is an $r \times r$ identity matrix, $\mathbf{0}$ is an $r \times s$ matrix with all 0's, $\mathbf{R}$ is an $s \times r$ matrix that specifies the transition from every transient state to every absorbing state, and $\mathbf{Q}$ is an $s \times s$ matrix that specifies the transition between the transient states. Note that $\mathbf{P}$ is row-stochastic, because the sum of each row is equal to 1, and $\boldsymbol{\pi}_t = \boldsymbol{\pi}_{t-1}\mathbf{P}$, where $\boldsymbol{\pi}_t$ is the state probability distribution vector at period t.

An input-output system can be modeled as an absorbing Markov chain, with the internal production stages as the transient states, and the external value-added contribution or final demand as the absorbing states. The process can be intuitively understood as a tour of a random walker. Depending on whether the Leontief model or the Ghosh model is used, the random walker is assigned the role of "purchasing agent" or "salesman." As stated above, we restrict our analysis to the Leontief model (i.e., the tour of a "purchasing agent" random walker).

As already mentioned, the total output $\mathbf{x}$ of a given sector can be written as the sum of the contributions of all n sectors and of the value-added $\mathbf{v}$, i.e., $\mathbf{1}^{\mathrm{T}}\mathbf{Z} + \mathbf{v}^{\mathrm{T}} = \mathbf{x}^{\mathrm{T}}$, where $\mathbf{1}^{\mathrm{T}}$ is a row-vector of 1's. Accordingly, a "purchasing agent" random walker which is in a given sector (a transient state) at period t can possibly visit other sectors (other transient states) at period $t+1$ or be absorbed by value-added (an absorbing state). To recast this dynamics into the canonical form (2), it is sufficient to specify $\mathbf{R}$ and $\mathbf{Q}$. In the tour of a "purchasing agent" random walker, the transient states are the n sectors in the country and the absorbing states are the n value-added suppliers corresponding to each sector. It turns out that the natural way to specify the transition matrix is to let $\mathbf{R} = \hat{\mathbf{v}}$ and $\mathbf{Q} = \mathbf{A}^{\mathrm{T}}$. Therefore, we have the following $2n \times 2n$ transition matrix (the subscript L denotes that it is based on the Leontief model):

$$\mathbf{P}_L = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \hat{\mathbf{v}} & \mathbf{A}^{\mathrm{T}} \end{bmatrix}, \tag{3}$$

where each submatrix has size $n \times n$. It is straightforward to check that $\mathbf{P}_L$ is row-stochastic.

Many theoretical results on absorbing Markov chains can fruitfully be exploited for input-output systems. For example, the *fundamental matrix* $\boldsymbol{\Lambda}$ of an absorbing Markov chain [31] is defined as

$$\boldsymbol{\Lambda} = \left(\mathbf{I} - \mathbf{A}^{\mathrm{T}}\right)^{-1} = \left[\left(\mathbf{I} - \mathbf{A}\right)^{-1}\right]^{\mathrm{T}} = \mathbf{L}^{\mathrm{T}}, \tag{4}$$

where the last equality is based on (1). The entry Λ_{ij} measures the average number of times that the "purchasing agent" random walker spends in sector j if the tour starts from sector i. Interestingly $\Lambda_{ij} = L_{ji}$, i.e., this quantity has a clear economic interpretation (see above).

2.2 Closed Input-Output Systems as Regular Markov Chains

In this paper, we model the input-output system as a *regular Markov chain*, i.e., without absorbing states. The easiest way to do that is to completely disregard the external value-added contribution and final demand, and to restrict to inter-industrial relationships only (see Fig. 1, left panels). This, however, implies a partial loss of information, that must be kept into account when discussing the results of the community analysis.

An alternative way to obtain a regular Markov chain is to endogenize the external information by "closing" the input-output system. This can be done by treating the final demand like another sector, called *household*, which buys products to consume and supplies production factors such as labor to sectors (i.e., from the view of a "salesman" random walker, it connects the material flow from final demand to value-added). On the other hand, the household receives payment for its supply of production factors, such as labor, and spends money to pay for the products from sectors (i.e., from the view of a "purchasing agent" random walker, it connects the monetary flow from value-added to final demand). Note that a more accurate definition of the household should actually be a mixture of households and governments. The latter is also an important player in both value-added and final demand (through fiscal policy). Value added is mostly supplied within national boundaries and the small share of foreign value added directly supplied in production cannot be estimated in a reliable way at the industry level, therefore we assume that value added is purely domestic. Final demand can be satisfied by foreign goods through imports, but it is conventional in input-output tables to treat imports as intermediates as they need the distribution sector to reach the consumers. Therefore we introduce one separate household sector for each country in the input-output system, linked domestically.

Fig. 1, right panels, shows a two-country, two-sector input-output system closed by two households, denoted by $H1$ and $H2$, respectively. Panel (c) is based on the view of material flow, while panel (d) is based on the view of monetary flow. Note that edges connected to the households are directed (one-way arrow) when they are cross-country: this is due to the constraint that a household can only supply value-added to its own country, which makes our model more realistic.

Along with the two ways mentioned above of converting the WION into a regular Markov chain, we therefore have four specifications of the network for community detection:

- *noH-SM*: no household, "salesman" random walker perspective (Fig. 1, panel (a)).
- *noH-PA*: no household, "purchasing agent" random walker perspective (Fig. 1, panel (b)).
- *H-SM*: with household, "salesman" random walker perspective (Fig. 1, panel (c)).

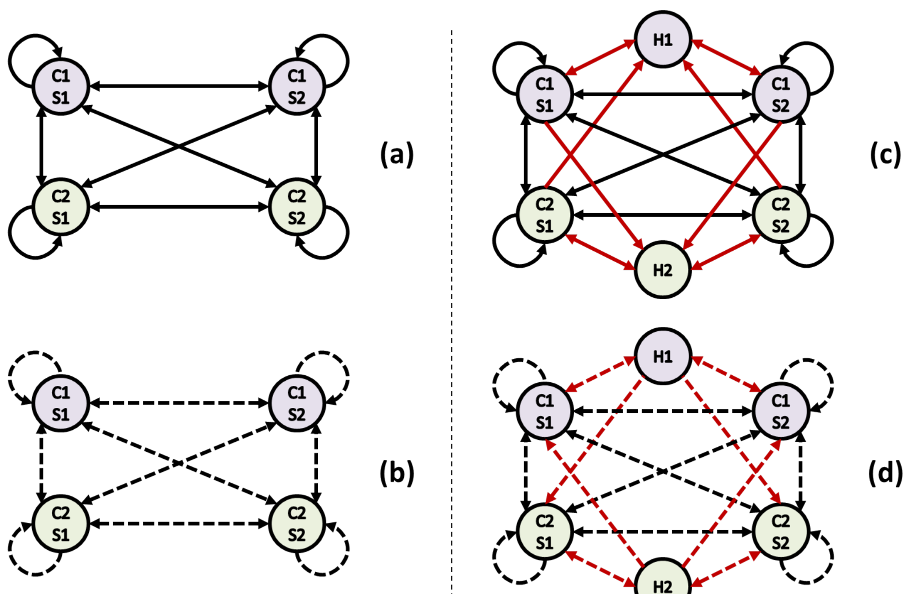

Figure 1: Closed input-output system without households (panels (a) and (b)) and with a household for each country (panels (c) and (d)) (all the edges related to the households are in red). Above: modeling the tour of a "salesman" random walker. Below: modeling the tour of a "purchasing agent" random walker.

- *H-PA*: with household, "purchasing agent" random walker perspective (Fig. 1, panel (d)).

2.3 Community Detection Based on Lumped Markov Chains

In the jargon of network analysis, a community is a subnetwork whose internal connections are comparatively stronger than the connections with the rest of the network [20, 21]. Finding communities in the WION, therefore, corresponds to highlighting sets of countries/sectors which are tightly connected, i.e., related by preferential economic relationships.

Among the plenty of community detection methods [20, 21], we use a Markov chain based method for consistency with the modeling approach of this paper, which is based on the equivalence between input-output models and random walks. Several Markov chain based community detection strategies have been proposed in the last decade (e.g., Walktrap [52], Infomap [57], Markov stability [14], and Linkrank [32]). In this paper, we use the Lumped Markov Chain approach [48], which allows the user to select the desired level of cohesiveness and provides a measure of the quality of each individual community. Furthermore, it allows one to assess the cohesiveness of exogenously defined subnetworks (e.g., countries or sectors), a feature that we will exploit in Sec. 3.3.

As described in Sec. 2.2, the dynamics of the random walker are modeled by a regular Markov chain. Given a network partition (i.e., a set of non-overlapping candidate communities), we will be able to describe the random walker dynamics by means of a "lumped" Markov chain, namely a simplified description where each state corresponds now to a community, rather than a node, and the transition probabilities among lumped nodes describe the motion of the random walker from one community to the other. Among the set of transition probabilities, we will focus our attention on the *Persistence Probabilities* (PP's), i.e., the probabilities that the random walker remains in the same community (or lumped node) where she is currently. Clearly, a large PP is the effect of a strong internal connectivity and is therefore indicative of a significant community. In this section, we succinctly review the methodology to make the paper self-contained – a detailed description can be found in [48].

Given a network with n nodes and weight matrix $\mathbf{W} = [W_{ij}]$, the transition probabilities of a random walker are standardly defined as $P_{ij} = W_{ij}/\sum_h W_{ih}$, i.e., the random walker in node i selects one of the out-edges with probability proportional to its weight. In our case, we use the transaction matrix $\mathbf{Z}$ as weight matrix (for the specifications *H-SM* and *H-PA* we also need information of $\mathbf{f}$ and $\mathbf{v}$). Note that using the P_{ij}'s corresponds to moving from absolute to relative trade values, since the flow $i \to j$ is now normalized by the total flow from sector i. The consequence is that communities will not necessarily be composed of groups of sectors related by large flows (in absolute terms) but, instead, of sectors with privileged partnership, namely, whose flows are important in relative terms. Let π_i be the probability of visiting node i in the long run or, equivalently, the fraction of time periods spent on node i in an infinitely long walk. In regular Markov chains, this quantity is well defined and strictly positive for all i [39]. Given a subnetwork S (e.g., a candidate community), it can be shown that its PP α_S, i.e., the probability that a random walker which is currently in any state of S remains in S at the next step, is given by

$$\alpha_S = \sum_{i \in S} \frac{\pi_i}{\Pi_S} \sum_{j \in S} \frac{W_{ij}}{\sum_h W_{ih}}, \tag{5}$$

where $\Pi_S = \sum_{i \in S} \pi_i$. This means that α_S is a weighted average (actually a convex combination) of the fractions of the outflows that the nodes (the country/sector pairs, in the WION case) of community S direct *within* the community itself. Large values of α_S are expected for significant communities, since the weights of the internal edges are comparatively large with respect to those pointing outside. For example, $\alpha_S = 0.5$ denotes that, on (weighted) average, the nodes of S direct half of their outflow within S and half to the rest of the network. The value $\alpha = 0.5$ can be used as a baseline threshold of significance, in line with the trade literature (see for example [13], [28], adopting the normalized average of trade value per country as a threshold). Moreover, it can easily be checked that, in the undirected/unweighed case, the constraint $\alpha \geq 0.5$ is equivalent to the well-established definition of "community" according to Radicchi et al. [54]. In Sec. 3.1, we will use the cutoff $\alpha = 0.5$ to show two detailed examples of our community detection and we will test a wide range of cutoff values in order to have robust findings.

The PP α_S allows one to assess the significance of a given, a priori defined, subnetwork S (e.g., the ensemble of sectors of a given country), to test whether it actually forms a cohesive set of nodes ("community testing", see Sec. 3.3 for the application to the WION case). But it is also a precious tool to detect communities endogenously, namely to derive a network partition composed of significant communities ("community detection", see Sec. 3.1). For that, we need a strategy for generating meaningful partitions, and we use with this aim a graph clustering procedure based, once again, on random walk dynamics.

We adopt a notion of similarity/distance among nodes analogous to [52, 61]. More precisely, we describe the global behavior of a large number of walkers (a "fleet") started from each node i, and we propose a similarity σ_{ij} between nodes (i, j) defined by

$$\sigma_{ij} = \sigma_{ji} = \frac{1}{T} \sum_{t=1}^{T} \left(\left[\mathbf{P}^t\right]_{ij} + \left[\mathbf{P}^t\right]_{ji} \right), \tag{6}$$

where $\mathbf{P} = [P_{ij}]$ is the transition matrix of the random walker. Then the distance $d_{ij} = d_{ji}$ between nodes (i, j) is defined by complementing the similarity and normalizing the results between 0 and 1:

$$d_{ij} = d_{ji} = 1 - \frac{\sigma_{ij} - \min \sigma_{ij}}{\max \sigma_{ij} - \min \sigma_{ij}}. \tag{7}$$

The rationale underlying the definition of σ_{ij} and d_{ij} is to assign nodes (i, j) a large similarity if a numerous fleet of random walkers started in i makes a large number of visits to j (and vice versa) within a sufficiently small time horizon T (see [48] for more technical details on the derivation of similarities). Nodes (i, j) are similar, therefore, if starting from i there is a large probability of visiting j within a few steps, and vice versa. Then a standard hierarchical, aggregative cluster analysis [17] is used to explore the possible existence of communities. More precisely, a binary cluster tree (dendrogram) is computed by initially defining n groups each containing a single node, and then by iteratively linking the two groups with minimal distance.

The community detection procedure is summarized in Fig. 2 for a 12-node toy network (panel (a)). The aggregative cluster analysis based on the distance (7) leads to a binary dendrogram (panel (b)). Horizontal top-down cross-sections of the dendrogram define a sequence of partitions composed of, respectively, $2, 3, \dots, q, \dots$ communities. For each partition, the PP's (5) of the q communities are computed (panel (c)) and, eventually, the selected partition is the one with the largest q provided all communities are cohesive above a given threshold, e.g., $\alpha_c \geq 0.5$ for all c. In other words, we take the finest decomposition (largest q) that guarantees a prescribed level of significance for all communities. From Fig. 2 we understand that if one sets a different (higher) cutoff value for α, then the number of communities reduces due to merging, since the entire procedure is based on hierarchical clustering. Tuning the threshold, therefore, corresponds to moving in the hierarchical tree towards finer or coarser partitions (see [9] for an early approach to hierarchically structured partitions).

In [48] the above methodology has extensively been tested on both real-world and benchmark networks (including different classes of LFR benchmarks [36, 35]). It proves to be more flexible and informative than max-modularity, as already pointed out, and it is more robust in identifying meaningful partitions when max-modularity suffers from scarce sensitivity (i.e., almost the same modularity value associated to many different partitions), notably when small communities have to be detected.

2.4 Community Centrality

Once a partition is obtained, we are interested in classifying nodes according to their importance within their own community, a property that we call *community centrality*. The economic relevance of a central position in an economic network has already been highlighted in the literature [27, 16]. Here we exploit the distance (7) above introduced and we straightforwardly extend the notion of closeness centrality [43] by attributing more importance to a node when, within a given community, it has smaller average distance to the other nodes of the same community. Formally, if we denote by $c(i)$ the community which node i belongs to, and by $n_{c(i)}$ its number of nodes, then the average distance from i to the nodes of $c(i)$ is

$$D_i = \frac{1}{n_{c(i)} - 1} \sum_{j \in \{c(i) \setminus i\}} d_{ij}, \tag{8}$$

and then the *community centrality* of node i can be defined as $\gamma_i = 1/D_i$.

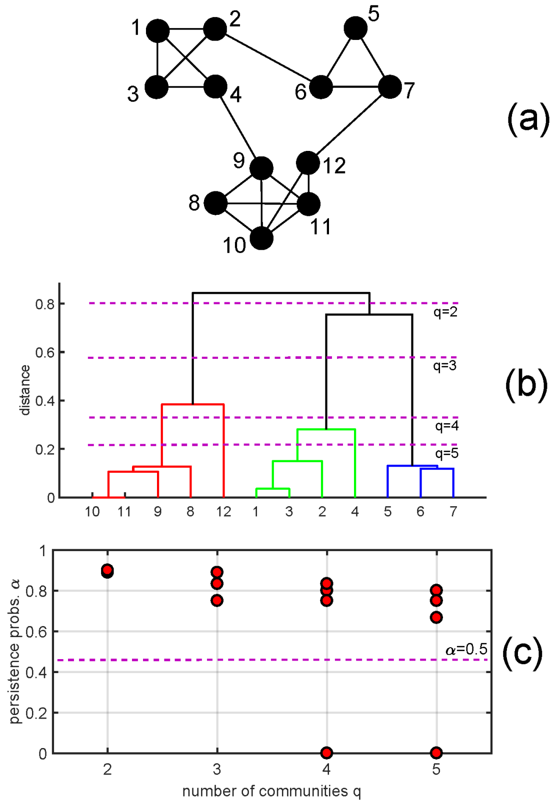

Figure 2: Summary of the community detection procedure. For the 12-node toy network of panel (a), subsequent cross-sections of the dendrogram of panel (b) define a set of partitions with $q = 2, 3, \ldots$ subnetworks. The finest decomposition (largest q) such that the PP's satisfy $\alpha_c \geq 0.5$ for all c is that with $q = 3$.

3 Results

We use the GMRIO tables of the World Input-Output Database [63], covering 35 sectors for each of the 40 countries (27 EU countries and 13 major countries in other regions) in the years from 1995 to 2011, to construct the World Input-Output Network (WION). The database provides yearly values of inputs sold and purchased by each sector in each country, the value added provided by factors of production, final demand in each country, and the total value of production. For each year, the nodes of the WION are the individual sectors across countries and the edges are the input-output relationships between them, which are directed and weighted. For all years and specifications, we checked that the WION is strongly connected: this is largely expected from the absence of source or sink nodes (no sector is self-sufficient, as it has to buy inputs and sell outputs) and has been empirically verified. Consequently, the stationary state distribution and the PP's (5) are well defined.

3.1 Community Detection with Different Specifications

We construct the WION for all the years and for the different network specifications listed at the end of Sec. 2.2. For the sake of brevity, here we only illustrate two cases in detail (all results are available by the authors upon request) while, in the next section, we will compare all cases in an aggregated form.

Fig. 3 shows the result of the community detection for the WION in 2011 with no households and "salesman" perspective, i.e., *noH-SM*. The community detection is conducted with the method described in Sec. 2.3 with a cutoff $\alpha = 0.5$ on PP's. In the figure, countries are arranged by rows and sectors are arranged by columns. Each cell represents a particular sector according to its column, in a particular country according to its row. Each color represents a community detected. For each community, numbers 1 to 3 denote the top three sectors according to community centrality. Rows are arranged in such a way that, from top to bottom, the communities decrease in size.

There are a number of interesting patterns emerging from Fig. 3. It is clear that most communities are essentially "row-based" (i.e., most sectors within each country belong to the same community) because of the strong domestic linkages. But we also observe some significant "column" communities (i.e., a significant number of countries within a specific sector belong to the same community), which indicate that these sectors are highly integrated across countries and are strongly affected by global value chains. For example, the Transport Equipment sector (#15) forms an inter-country cluster merged with the inter-sector community involving Germany, while Textiles (#4) and Leather (#5) across many countries are clustered into the community involving Italy. It is worth mentioning that we find some large communities such as the one integrating Germany with a number of Eastern Europe countries (in red) and the Asia-Pacific community (in purple). If, for each community, we mark the top three nodes in terms of community centrality, we find that the central sectors mostly include Construction (#18) and Business services (#30), but primary

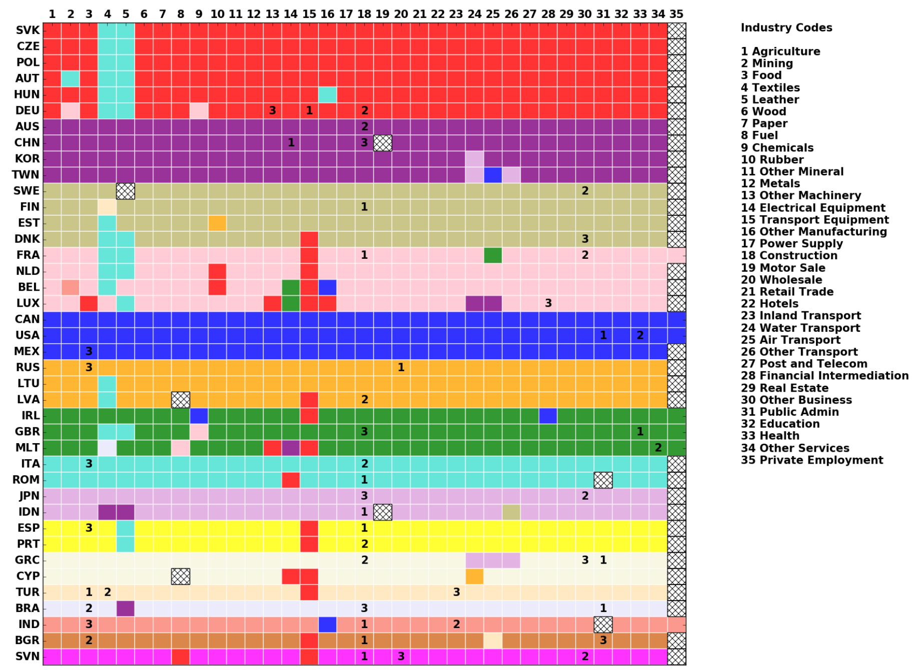

Figure 3: Communities in the WION, case *noH-SM*, year 2011: each color represents a community. For each community, numbers 1 to 3 denote the top three sectors according to community centrality. Texture indicates unavailable data.

sectors such as Food (#3) are more likely to be central for developing countries.

A complementary case is that in Fig. 4, which shows the communities detected with households and "purchasing agent" perspective, i.e., *H-PA* in 2011, again with a cutoff $\alpha = 0.5$ on PP's. A number of differences stand out immediately. There are more communities found in the case *H-PA* (32 communities) than in the case *noH-SM* (16 communities). This is largely due to the inclusion of the extra household sector (#36), which typically has strong domestic linkages and makes the communities more "country-based" (i.e., each row forms a community). Moreover, the "row" and "column" communities are very different once we switch from the "salesman" perspective to the "purchasing agent" perspective. For example, the large community led by Germany in Fig. 3 disappears in Fig. 4 where Germany only integrates with Austria. Also, Italy integrates with Romania and with Malta respectively in Fig. 3 and in Fig. 4. Regarding the "column" communities, we only find that the Fuel sector (#8) across a few countries is clustered into the community involving Russia. These differences highlight the fact that the upstream relationships ("purchasing agent" perspective) are hardly symmetric with the downstream ones ("salesman" perspective). Finally, in the *H-PA* case the household sector (#36) turns out to be very central. This is not surprising since the household sector endogenizes the external value-added and final demand, which take a large portion of the total output for each country.

In the two cases above discussed, the standard cutoff value $\alpha = 0.5$ on PP's has been used. In Fig. 5 we track how the number of detected communities (averaged over the time interval spanned by our analysis) depends on the cutoff value α. Note, first of all, that fewer and fewer communities are systematically detected when α increases, regardless to the modeling specification, for the reasons discussed in Sec. 2.3. The cases with households *H-SM* and *H-PA* have more communities than those with no households *noH-SM* and *noH-SM*, for a wide range of α values. As a matter of fact, once the external information of value-added and final demand is kept by the households, less integration is observed between countries because the households introduce strong linkages within each individual country, which becomes more likely to be detected as a single cohesive community. Recall that, by our definition, the households only supply value-added domestically and the value-added share of a sector's total output is typically large.

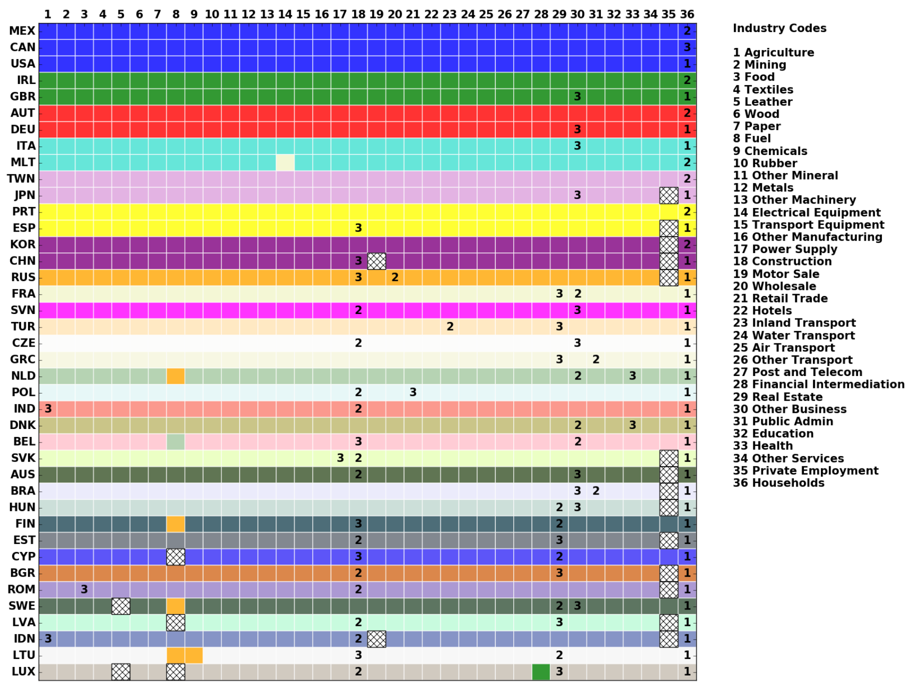

Figure 4: Communities in the WION, case *H-PA*, year 2011: each color represents a community. For each community, numbers 1 to 3 denote the top three sectors according to community centrality. Texture indicates unavailable data.

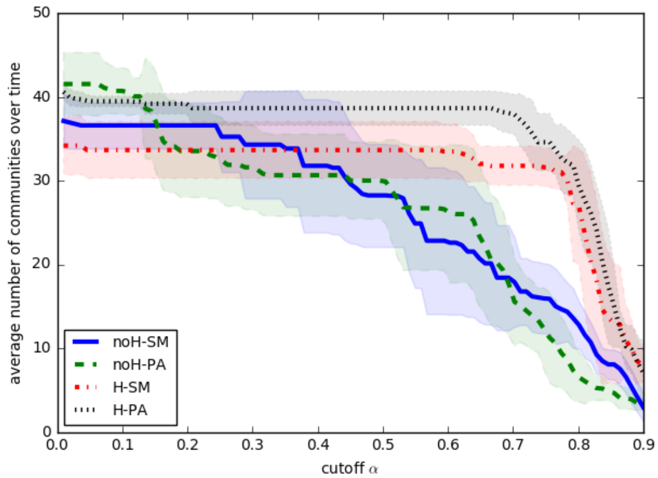

Figure 5: The impact of the cutoff value α on the results of community analysis, measured by the number of detected communities for the four modeling specifications. For each value of the cutoff α (100 data points from 0.01 to 0.9), the average ($\pm$ std) number of communities over time is plotted.

In order to further assess the sensitivity to the cutoff value α, we compare our results, in aggregated form, to those obtained by max-modularity, the most popular community detection approach [20, 21]. To compare two partitions of the same network, we use the Normalized Variation of Information (NVI) [38], which is defined as

$$NVI = -\frac{1}{\log n} \sum_{c} \sum_{c'} \frac{n_{cc'}}{n} \log \frac{(n_{cc'})^2}{n_c n_{c'}}, \tag{9}$$

where c, c' iterate, respectively, through the set of communities of the two partitions, $n_{cc'}$ is the number of common nodes in communities c and c', n_c is the number of nodes in community c, and n is the total number of nodes of the network (conventionally we set $0 \log 0 = 0$). The quantity NVI varies between 0 and 1, the lower bound 0 being attained when the two partitions are identical, and the upper bound 1 when they are maximally different (i.e., one partition has n clusters and the other only 1).

Fig. 6 shows the results of the comparison, for the four modeling specifications described in Sec. 2.2. The two methods do produce different results, with an average NVI value around 0.3. However, the magnitude of the NVI appears to be independent from α except for very high values. This implies that the information captured by our method is partially different from that captured by max-modularity (we refer to [48] for a thorough discussion on this point) and, most important, that the difference is robust against a wide range of α.

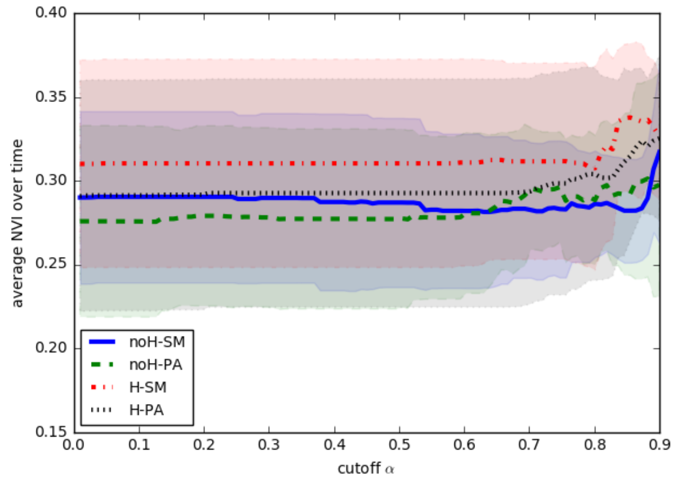

Figure 6: The impact of the cutoff value α on the results of community analysis, measured by the Normalized Variation of Information (NVI) between the partition resulting from our community detection and that obtained via max-modularity, for the four modeling specifications. For each value of the cutoff α (100 data points from 0.01 to 0.9), the average ($\pm$ std) NVI over time is plotted.

3.2 Community Dynamics

We perform a preliminary analysis, in aggregated form, of the time evolution of the communities detected with the method above described. We note that a thorough analysis of the mesoscale dynamics requires the use of systematic methods (e.g., [42, 2, 47]) which, however, would imply to enter into theoretical and methodological aspects that cannot be included in this work, and will instead be considered in a future research. Here we simply highlight some evidence emerging when comparing the different modeling specifications.

For each cutoff value α, we first compute the NVI between the partitions obtained in two consecutive years (16 pairs, from 1995-1996 to 2010-2011) for any given modeling specification, and then compute the average NVI over time. The results are in Fig. 7, which shows that the NVI values are weakly sensitive to the cutoff value α.

Overall, the general low value of the NVI index suggests that emerging communities are quite stable over time, confirming that the economic structure and the production linkages are not changing quickly, due to factors such as technological constraints and trade policy arrangements. Within this narrow range, the partitions including households (*H-PA* and *H-SM*) are generally even more stable than the others, suggesting that final demand and use of factors of production have especially persistent links, while production links (*noH* partitions) display more changes, as firms look for better opportunities in terms of markets where to sell and where to purchase inputs. This is consistent with our previous finding that the inclusion of households creates strong domestic linkages and makes the communities more "country-based" and more robust over time. We also observe that the "salesman" perspective shows more volatility than the "purchasing agent" perspective, and this is true with or without households. As a matter of fact, downstream relationships are more open (i.e., internationally distributed) than upstream ones. Therefore, the former is more exposed to uncertainty and risk and is more volatile than the latter. As showcased in Figs. 3 and 4, the "salesman" perspective has more economic integration across country borders than the "purchasing agent" perspective.

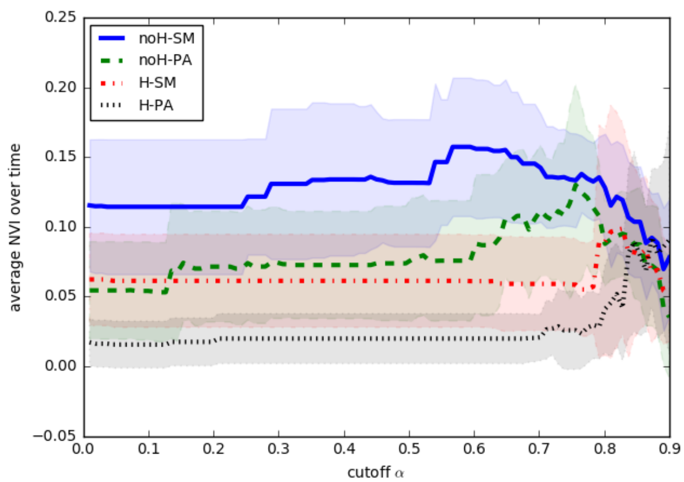

Figure 7: Community dynamics: for each cutoff value α (100 data points from 0.01 to 0.9), we plot the average ($\pm$ std) NVI value over the 16 pairs of consecutive years (from 1995-1996 to 2010-2011), for the four modeling specifications.

3.3 Evaluation of Exogenously Defined Communities

The methodology for community analysis adopted in this work [48] enables one, as already pointed out, to assess the cohesiveness of any predefined subnetwork communities. This is especially useful in the WION context because, for instance, we can track the PP over time for a predefined set of nodes corresponding to all sectors in a given country or, dually, to all countries for a given sector. Moreover, since the PP relates to the fraction of trade internally directed by the nodes of the subnetwork (see Sec. 2.3), for a sector community, unless its

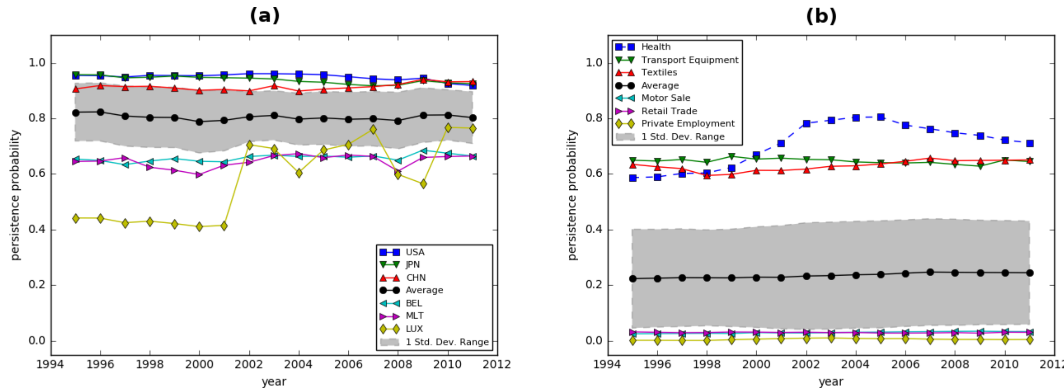

Figure 8: The evolution in time of the persistence probabilities of predefined country communities (panel (a)) and sector communities (panel (b)), case *noH-SM*. For both panels, the average PP (±std) and the top and bottom three communities in terms of their average PP over time are shown. Note that the health sector in panel (b) is shown with a dashed line because its high PP value is not due to high "openness" but rather to strong self-loops.

nodes are characterized by strong self-loops (i.e., a significant amount of inputs going from one sector to itself), we can interpret its PP as a direct measure of the "openness," i.e., tendency to globalization, because a more cohesive sector community indicates stronger edges between countries. On the other hand, for a country community, we can interpret its PP's complement to one (i.e., one minus the PP value) as a measure of "openness" because a more cohesive country community indicates weaker edges between countries.

We focus on the specification *noH-SM* and track the PP for the predefined country and sector communities for all the years from 1995-2011. The evolution of the PP's for the predefined country communities is shown in panel (a) of Fig. 8. In particular, we show the average PP across all countries (with its one-standard-deviation range) and the top and bottom three countries in terms of their average PP over time. The average PP stays around 0.8 over time. Therefore, its complement to one is only about 0.2 and indicates that countries are relatively closed. This is actually expected, and proves that most of input-output transactions still happen within country borders: this is confirmed by the dominating "row" communities detected in Fig. 3. However, there is heterogeneity behind this average pattern: small countries such as Luxembourg, Malta, and Belgium tend to have relatively high openness, while big countries such as the USA, China, and Japan tend to have very low openness. This again fits the economic intuition very well, because small countries are more dependent on international resources and markets than big ones.

Panel (b) of Fig. 8 shows the PP's of the predefined sector communities over time. The average PP is stable over time with its value slightly above 0.2 and indicates that sectors are also relatively closed and the cross-country linkages within the same sector are typically weak. However, the sector communities exhibit greater heterogeneity than the country ones in panel (a). First, compared with the situation in panel (a), the values of PP's are more dispersed for the sector communities as the one-standard-deviation range is much wider. Second, sectors like health, transport equipment, and textiles appear to be "outliers", being more than two standard deviations away from the average. Nevertheless, as stated above, whether the high PP values imply more globalization requires further investigation. We find out that the high value of the health sector is driven by strong self-loops. That is, the health sector in each country typically takes a dominant amount of inputs from itself. As an important service sector, it is in fact very closed (therefore a dashed line is used to differentiate it from others). In contrast, the sectors of transport equipment and textiles are indeed globalized as they are revealed in Fig. 3 as the significant "column" communities.

4 Conclusions

In this paper we use the GMRIO tables from the World Input-Output Database [63] to construct the World Input-Output Network (WION), where the nodes are the individual sectors across countries and the edges are the input-output relationships between them. We first convert the open world input-output system (an absorbing Markov chain) to a regular Markov chain by either removing the external value-added contribution and final demand, or by endogenizing them. Then we determine the meaningful partitions of the WION by transforming the regular Markov chain to a lumped Markov chain and evaluating the *persistence probabilities* of each lumped community. The random walks on the WION can be implemented either from the perspective of "salesman" random walker (supply side) or from the perspective of a "purchasing agent" random walker (demand side). Along with the above two ways of converting the WION to a regular Markov chain, we have in total four specifications to conduct community detection.

The main contribution of this paper is precisely the application of techniques of community detection to GMRIO tables, allowing to highlight the deep international connections existing between production systems of different countries. This procedure also suggests a different method to measure countries participation in global value chains. In addition, by determining which nodes (which sectors in which countries) are most central in the resulting communities, it underscores the industries playing an important role in the world economic structure.

There exist many community detection methods and the choice is always context-dependent. The random walk approach proves more suitable in the context of the WION when compared to traditional community detection methods such as max-modularity. First, it can naturally be applied to both national and international input-output systems, because the latter can be shown to be closely related to Markov chains. Second, rather than providing only a partition and a global network measure, as max-modularity does, it quantifies the significance of each detected community via persistence probabilities. Third, it allows the evaluation of the cohesiveness of any subnetwork exogenously defined.

Many interesting economic interpretations can be drawn from the empirical results. First, the two specifications without household show richer dynamics than the two with household. This is the footprint of a strongly connected and constantly evolving world production system. The difference obtained when including or excluding the household sector also underscores the key role played by final demand and factors of production in an economic system. Second, although both *noH-SM* and *noH-PA* show higher integration level than the other specifications, the community structure with *noH-SM* is more volatile over time than that with *noH-PA*. This difference is in line with the suggestion of the economic literature that it matters whether participation to international production processes is observed in terms of upstream or downstream linkages [34], as the choice of suppliers and destination markets of firms might follow different criteria. Finally, by evaluating the quality of predefined country and sector communities, we find evidence of heterogeneous globalization levels among countries and sectors.

For future work, we plan to examine the impact of the position of a given sector in the WION on its macroeconomic performance, and to test whether the comovement (synchronization) between sectors in terms of macroeconomic dynamics can be explained by their proximity in the WION. Another topic worth exploring is to simulate the introduction in the WION of some policy changes (such as trade tariffs and business taxes) and analyze how the system evolves to a new equilibrium structure.

Funding

The work of M.R. and Z.Z. was supported by the "ViWaN: The Global Virtual Water Network" project and the National Research Project (PNR) CRISIS Lab.

References

[1] Acemoglu, D., Carvalho, V. M., Ozdaglar, A. & Tahbaz-Salehi, A. (2012) The network origins of aggregate fluctuations. *Econometrica*, **80**(5), 1977–2016.

[2] Bassett, D. S., Porter, M. A., Wymbs, N. F., Grafton, S. T., Carlson, J. M. & Mucha, P. J. (2013) Robust detection of dynamic community structure in networks. *Chaos*, **23**(1).

[3] Battiston, S., Puliga, M., Kaushik, R., Tasca, P. & Caldarelli, G. (2012) Debtrank: Too central to fail? financial networks, the fed and systemic risk. *Scientific reports*, **2**, 541.

[4] Carvalho, V. M. (2014) From micro to macro via production networks. *The Journal of Economic Perspectives*, **28**(4), 23–47.

[5] Cerina, F., Zhu, Z., Chessa, A. & Riccaboni, M. (2015) World Input-Output Network. *PLoS ONE*, **10**(7), 1–21.

[6] Chan-Lau, J. A. (2010) Regulatory Capital Charges for Too-Connected-to-Fail Institutions: A Practical Proposal. *Financial Markets, Institutions & Instruments*, **19**(5), 355–379.

[7] Chessa, A., Morescalchi, A., Pammolli, F., Penner, O., Petersen, A. M. & Riccaboni, M. (2013) Is Europe evolving toward an integrated research area?. *Science*, **339**(6120), 650–651.

[8] Cingolani, I., Piccardi, C. & Tajoli, L. (2015) Discovering Preferential Patterns in Sectoral Trade Networks. *PLoS ONE*, **10**, 1–16.

[9] Clauset, A., Moore, C. & Newman, M. E. J. (2008) Hierarchical structure and the prediction of missing links in networks. *Nature*, **453**(7191), 98–101.

[10] Contreras, M. G. A. & Fagiolo, G. (2014) Propagation of economic shocks in input-output networks: A cross-country analysis. *Physical Review E*, **90**(6), 062812.

[11] De Benedictis, L., Nenci, S., Santoni, G., Tajoli, L. & Vicarelli, C. (2014) Network Analysis of World Trade using the BACI-CEPII dataset. *Global Economy Journal*, **14**(3-4), 287–343.

[12] De Benedictis, L. & Tajoli, L. (2011) The world trade network. *The World Economy*, **34**(8), 1417–1454.

[13] De Lombaerde, P. & Iapadre, L. (2008) The world is not flat: implications for the construction of globalization indicators. *World Economics*, **9**(4), 159–180.

[14] Delvenne, J. C., Yaliraki, S. N. & Barahona, M. (2010) Stability of graph communities across time scales. *Proceedings of the National Academy of Sciences of the United States of America*, **107**(29), 12755–12760.

[15] Duchin, F. & Levine, S. H. (2010) Embodied resource flows and product flows. *Journal of Industrial Ecology*, **14**(4), 586–597.

[16] Duernecker, G., Meyer, M. & Vega-Redondo, F. (2014) The network origins of economic growth, `http://nbn-resolving.de/urn:nbn:de:bsz:180-madoc-354839`. *Working Paper Series, Department of Economics, University of Mannheim, No. 14-06.*

[17] Everitt, B. S., Landau, S., Leese, M. & Stahl, D. (2011) *Cluster Analysis, 5th ed.* John Wiley & Sons.

[18] Fagiolo, G., Reyes, J. & Schiavo, S. (2009) World-trade web: Topological properties, dynamics, and evolution. *Physical Review E*, **79**(3), 036115.

[19] Ferrier, G., Reyes, J. & Zhu, Z. (2016) Technology Diffusion on the International Trade Network. *Journal of Public Economic Theory*, **18**(2), 291–312.

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

[21] Fortunato, S. & Hric, D. (2016) Community detection in networks: A user guide. *Physics Reports*, **659**, 1–44.

[22] Foschi, R., Riccaboni, M. & Schiavo, S. (2014) Preferential attachment in multiple trade networks. *Physical Review E*, **90**(2), 022817.

[23] Fronczak, A. & Fronczak, P. (2012) Statistical mechanics of the international trade network. *Physical Review E*, **85**(5), 056113.

[24] He, J. & Deem, M. W. (2010) Structure and response in the world trade network. *Physical Review Letters*, **105**(19), 198701.

[25] Hertel, T. (1997) *Global Trade Analysis: modeling and applications*. Cambridge University Press.

[26] Higashi, M., Patten, B. C. & Burns, T. P. (1993) Network trophic dynamics: the modes of energy utilization in ecosystems. *Ecological modelling*, **66**(1-2), 1–42.

[27] Hochberg, Y., Ljungqvist, A. & Lu, Y. (2007) Whom you know matters: venture capital networks and investment performance. *Journal of Finance*, **62**(1), 251–301.

[28] Hoekman, B., Martin, W. & Primo Braga, C. (2009) *Trade Preference Erosion: Measurement and Policy Response*. Palgrave Macmillan.

[29] Johnson, R. & Noguera, G. (2012) Accounting for Intermediates: Production Sharing and Trade in Value Added. *Journal of International Economics*, **86**(2), 224–236.

[30] Kelly, G., Cooper, A. & Pinkerton, E. (2014) Social network analysis, Markov chains and input-output models: Combining tools to map and measure the circulation of currency in small economies. *Journal of Rural and Community Development*, **9**(3), 118–141.

[31] Kemeny, J. & Snell, J. (1983) *Finite Markov Chains: With a New Appendix "Generalization of a Fundamental Matrix"*. Undergraduate Texts in Mathematics. Springer New York.

[32] Kim, Y., Son, S.-W. & Jeong, H. (2010) Finding communities in directed networks. *Physical Review E*, **81**(1, Part 2), 016103.

[33] Kitsak, M., Riccaboni, M., Havlin, S., Pammolli, F. & Stanley, H. E. (2010) Scale-free models for the structure of business firm networks. *Physical Review E*, **81**(3), 036117.

[34] Koopman, R., Wang, Z. & Wei, S. (2014) Tracing Value-Added and Double Counting in Gross Exports. *American Economic Review*, **104**(2), 459–494.

[35] Lancichinetti, A. & Fortunato, S. (2009) Benchmarks for testing community detection algorithms on directed and weighted graphs with overlapping communities. *Physical Review E*, **80**(1, Part 2), 016118.

[36] Lancichinetti, A., Fortunato, S. & Radicchi, F. (2008) Benchmark graphs for testing community detection algorithms. *Physical Review E*, **78**(4, Part 2), 046110.

[37] Leontief, W. (1954) Domestic Production and Foreign Trade: The American Capital Position Reexamined. *Economia Internazionale*, **7**, 3–32.

[38] Meilă, M. (2007) Comparing clusterings: an information based distance. *Journal of multivariate analysis*, **98**(5), 873–895.

[39] Meyer, C. (2000) *Matrix Analysis and Applied Linear Algebra*. SIAM.

[40] Miller, R. & Blair, P. (2009) *Input-Output Analysis: Foundations and Extensions*. Cambridge University Press.

[41] Morrison, G., Giovanis, E., Pammolli, F. & Riccaboni, M. (2014) Border sensitive centrality in global patent citation networks. *Journal of Complex Networks*, **2**(4), 518–536.

[42] Mucha, P. J., Richardson, T., Macon, K., Porter, M. A. & Onnela, J.-P. (2010) Community structure in time-dependent, multiscale, and multiplex networks. *Science*, **328**(5980), 876–878.

[43] Newman, M. (2010) *Networks: an introduction*. Oxford University Press.

[44] Newman, M. E. & Girvan, M. (2004) Finding and evaluating community structure in networks. *Physical Review E*, **69**(2), 026113.

[45] Onnela, J., Chakraborti, A., Kaski, K., Kertesz, J. & Kanto, A. (2003) Dynamics of market correlations: Taxonomy and portfolio analysis. *Physical Review E*, **68**(5, 2).

[46] Pammolli, F. & Riccaboni, M. (2002) Technological regimes and the growth of networks: An empirical analysis. *Small Business Economics*, **19**(3), 205–215.

[47] Peixoto, T. P. & Rosvall, M. Modeling sequences and temporal networks with dynamic community structures, `https://arxiv.org/abs/1509.04740v2`. .

[48] Piccardi, C. (2011) Finding and testing network communities by lumped Markov Chains. *PLoS ONE*, **6**(11), 1–13.

[49] Piccardi, C., Calatroni, L. & Bertoni, F. (2010) Communities in Italian corporate networks. *Physica A: Statistical Mechanics and its Applications*, **389**, 5247–5258.

[50] Piccardi, C., Calatroni, L. & Bertoni, F. (2011) Clustering financial time series by network community analysis. *International Journal of Modern Physics C*, **22**(1), 35–50.

[51] Piccardi, C. & Tajoli, L. (2012) Existence and significance of communities in the world trade web. *Physical Review E*, **85**(6), 066119.

[52] Pons, P. & Latapy, M. (2005) Computing communities in large networks using random walks. In Yolum, P. and Gungor, T. and Gurgen, F. and Ozturan, C., editor, *Computer and Information Sciences - ISCIS 2005, Proceedings*, volume 3733 of *Lecture Notes In Computer Science*, pages 284–293. Springer-Verlag Berlin.

[53] Powell, W. W., Koput, K. W. & Smith-Doerr, L. (1996) Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology. *Administrative Science Quarterly*, **41**, 116–145.

[54] Radicchi, F., Castellano, C., Cecconi, F., Loreto, V. & Parisi, D. (2004) Defining and identifying communities in networks. *Proceedings of the National Academy of Sciences of the United States of America*, **101**(9), 2658–2663.

[55] Riccaboni, M. & Pammolli, F. (2002) On firm growth in networks. *Research Policy*, **31**(8), 1405–1416.

[56] Riccaboni, M. & Schiavo, S. (2014) Stochastic trade networks. *Journal of Complex Networks*, **2**(4), 537–556.

[57] Rosvall, M. & Bergstrom, C. T. (2008) Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences of the United States of America*, **105**(4), 1118–1123.

[58] Rzhetsky, A., Foster, J. G., Foster, I. T. & Evans, J. A. (2015) Choosing experiments to accelerate collective discovery. *Proceedings of the National Academy of Sciences of the United States of America*, **112**(47), 14569–14574.

[59] Serrano, M. A. & Boguná, M. (2003) Topology of the world trade web. *Physical Review E*, **68**(1), 015101.

[60] Sgrignoli, P., Metulini, R., Schiavo, S. & Riccaboni, M. (2015) The relation between global migration and trade networks. *Physica A: Statistical Mechanics and its Applications*, **417**, 245–260.

[61] Steinhaeuser, K. & Chawla, N. V. (2010) Identifying and evaluating community structure in complex networks. *Pattern Recognition Letters*, **31**(5), 413–421.

[62] Suh, S. (2005) Theory of materials and energy flow analysis in ecology and economics. *Ecological modelling*, **189**(3), 251–269.

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

[64] Tzekina, I., Danthi, K. & Rockmore, D. N. (2008) Evolution of community structure in the world trade web. *European Physical Journal B*, **63**(4), 541–545.

[65] Wiedmann, T., Wilting, H. C., Lenzen, M., Lutter, S. & Palm, V. (2011) Quo Vadis MRIO? Methodological, data and institutional requirements for multi-region input-output analysis. *Ecological Economics*, **70**(11), 1937 – 1945. Special Section - Earth System Governance: Accountability and Legitimacy.

[66] Zhu, Z., Cerina, F., Chessa, A., Caldarelli, G. & Riccaboni, M. (2014) The Rise of China in the International Trade Network: A Community Core Detection Approach. *PLoS ONE*, **9**(8), 1–8.

[67] Zhu, Z., Morrison, G., Puliga, M., Chessa, A. & Riccaboni, M. The Similarity of Global Value Chains: A Network-Based Measure, `https://arxiv.org/abs/1508.04392`. .

[68] Zhu, Z., Puliga, M., Cerina, F., Chessa, A. & Riccaboni, M. (2015) Global Value Trees. *PLoS ONE*, **10**(5), 1–17.