Physica A 527 (2019) 121178

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Input–output networks offer new insights of economic structure

Ming Xu [a,b], Sai Liang [c,*]

[a] School for Environment and Sustainability, University of Michigan, Ann Arbor, MI 48109-1041, United States
[b] Department of Civil and Environmental Engineering, University of Michigan, Ann Arbor, MI 48109-2125, United States
[c] State Key Joint Laboratory of Environment Simulation and Pollution Control, School of Environment, Beijing Normal University, Beijing, 100875, People's Republic of China

HIGHLIGHTS

- We introduce IO network analysis framework including several popular metrics and tools.
- This study provides additional insights for better understanding economic structures.
- We show the framework using hypothesized and real-world demonstrations.

ARTICLE INFO

Article history:
Received 14 January 2019
Received in revised form 9 April 2019
Available online 13 May 2019

Keywords:
Input–output model
Network analysis
Economic structure

ABSTRACT

An input–output (IO) model can be regarded as a network in which nodes represent sectors and directional, weighted links stand for IO transactions between sectors. The integration of IO models with modern network analysis can potentially provide additional insights for better understanding the structure of economies. We introduce the framework of IO network analysis including several popular metrics and tools. We also demonstrate the framework using a hypothesized six-sector economy. The World Input–Output Database (WIOD) 2009 model is used as well for a real-world demonstration. This research shows the potential of IO network analysis in understanding the structure of economies using IO models and data. Our work lays the ground for future studies in developing new methods for IO network analysis and real-world case studies.

1. Introduction

An economy comprises sectors that are interdependent with each other through the exchange of products and services. Understanding the structure of the economy – how sectors are interconnected with each other – is important for understanding how the system works, because the structure of a system usually determines its functionality. The input–output (IO) model characterizes the structure of an economy at the sector level [1–3] and the interaction between the economy and the environment through flows of energy, resources, and emissions [4–13]. The most important feature of the IO model is that it tracks both direct and indirect supply–demand interdependencies among sectors within the entire economy. This allows understanding the relationships between components of a system without putting an individual component under microscope in isolation from other components that directly or indirectly interact with it.

Classic IO analysis packs all inter-sectoral relationships together as a technical coefficient matrix **A**. The **A** matrix and the consequent Leontief inverse matrix **L** provide great computational convenience to measure the direct and indirect

* Corresponding author.
E-mail address: liangsai@bnu.edu.cn (S. Liang).

https://doi.org/10.1016/j.physa.2019.121178

interdependencies between sectors in an economy. However, this great convenience may also come with a great cost. Packaging the rich information on the complex sector-level interactions all together as a matrix almost like a black-box greatly limits our ability of investigating the structural features of the economy at the sector level. What is the role of a particular sector or particular sectors in the economy? What does it mean to the entire economy for particular sectors interact with each other in certain ways? What is the relationship between the performance of the economy and the way how sectors are interconnected with each other? These are important policy-relevant questions, answering which requires characterizing and understanding the structure of the economy at the sector level. The traditionally highly packaged rich information on sectoral interdependencies from IO models represents unprecedented opportunity for a much deeper understanding of how sectors connect, interact, and interdepend with each other (or the "interconnectedness" of sectors) and how the structure of the economy at the sector level determines the functionality of the entire economy.

Existing tools based on IO models touch on some aspects of understanding the structural features of an economy, such as structural path analysis (SPA) and linkage analysis [14–17]. However, these tools still reflect somewhat of a reductionism in which components of a system are still studied in isolation from other components. For example, SPA is interested in measuring contributions of *separate* paths to *particular* sectors. Linkage analysis moves one step further by studying the upstream and downstream impacts to the entire economy, but focuses on the impacts due to changes in, still, *particular* sectors. Little attention in the IO literature has been paid to study the structure of an economy as a complete, integrated system. Just as we cannot comprehend the dynamics of ecosystems by studying separated food chains [18] or the behavior of cells by micro-scoping isolated biochemical pathways [19,20], we cannot fully understand the structure and its relation to the dynamics of an economy only by investigating isolated, separated interdependencies between sectors. A more "holistic" perspective is required. Modern "network analysis" [21–24] offers an ideal framework to pursue such as study using IO models.

Rooted in graph theory, modern network analysis emerges as a data-driven approach for characterizing the structure of large-scale network-like complex systems to infer the causality between the structure and functionality of the system [25]. Those systems often studied in the network analysis literature span many disciplines, such as social networks [26], biological and metabolic networks [20,27], food chain networks in ecosystems [28], and transportation networks [29]. In particular, a network consists of nodes (or vertexes) that are connected with each other through links (or edges). The structure of a network essentially reflects the specific way that nodes are connected by links. The structure of those real-world networks is found often to be extremely heterogeneous in the way that a few nodes are disproportionately highly connected while the majority of nodes are loosely connected, formally known as the power-law distribution [30–33]. Note that nodes and links could also be heterogeneous themselves in many ways (e.g., size, strength). Such heterogeneity indicates components of a system play different roles in the system, thus implies that the structure of the system is critical to the functionality of the system. Network analysis provides a suite of methods and metrics to comprehend the structure of a network and relate it to the functionality of the system represented by the network.

An IO-based economy can be regarded as a network in which nodes represent sectors and links connecting nodes stand for the economic transactions between sectors in the IO model. In Fig. 1, we plot the complementary cumulative density functions (CCDFs) of the node weights (outputs of sectors) and link weights (value of inter-sector transactions) for four IO networks based on four IO models: World Input–Output Database (WIOD) [34] and Eora [35] for the world economy in 2009 and 2010, respectively, Bureau of Economic Analysis (BEA) 2002 (www.bea.gov) for the US economy, and a Chinese multi-regional input–output (MRIO) model for 2007 [36] at the provincial level. Power-law distributions or similar ones are observed for these IO networks, indicating heterogeneity structure of IO networks and the potential suitability of using network analysis to provide insights for understanding the effects of structural features on the economy.

The need for studying the structure of an economy as a network was proposed in as early as 1978. Slater [37] applies a maximum flow minimum cut algorithm from network theory to the 1967 US IO table to identify "production and consumption complexes" which essentially mean sector clusters in the IO network. However, the literature went in silence until about three years ago. Recent studies use national IO tables to characterize the structure of respective economies, mainly focusing on three areas. First, the positions of particular sectors in the IO network are examined using a variety of centrality measures such as betweenness [10,38]. The underlying premise is that sectors that are closer to the center of the network than others tend to be more important to the economy. Notably, a subset of this literature applies ecological network analysis – a method that was used to study structural features of ecological systems – to IO data with environmental extensions to understand how the structure of an economy affects environmental profile of nations [39–43].

Second, researchers apply various methods from network analysis to identify clusters or communities in IO networks. Clusters essentially are sectors within the IO network that are strongly connected with each other as hubs. The strength of the connections between sectors are measured by either the value of goods and services exchanged [44] or the amount of emissions associated with such exchanges [45,46]. Lastly, the structure of IO networks is related to the performance of economies they represent, primarily on the resilience of the economy against economic and information shocks [10,38,47,48], but also the development of economies [49,50].

How to characterize the structure of the IO network alone is an interesting research avenue, because it represents the most complex network possible – a directional, node- and link-weighted, and self-connected network – and IO networks are usually significantly denser than previous studied networks (e.g., hundreds of links per node in IO networks vs. several or dozens in other networks). The current literature primarily focuses on analyzing the mathematical and statistical

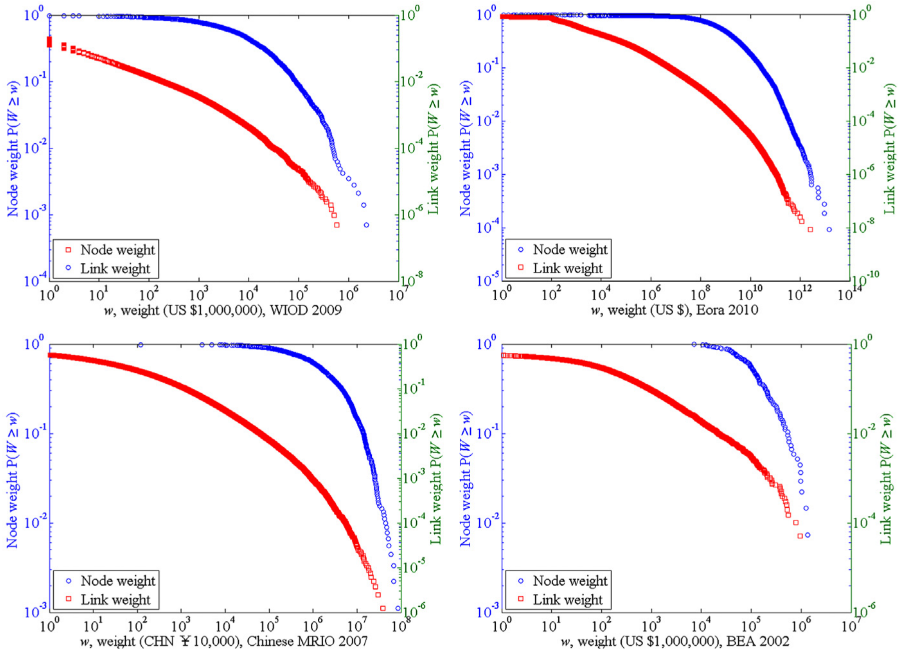


Fig. 1. Complementary cumulative density functions (CCDFs) of node weights (sectoral outputs) and link weights (inter-sectoral transaction values) for selected input–output (IO) networks. P(W $\geq$ w) measures the probability that the weight of a node (or link) in the network is equal to or larger than a certain value (w). The upper left Figure is for the network based on the 2009 data from World Input-Output Database (WIOD, www.wiod.org) The upper right Figure is for the network based on the 2010 data from Eora multi-regional input–output (MRIO) database (worldmrio.com). The lower left Figure is for the network based on Chinese 2007 MRIO table from Liu et al. [36]. The lower right Figure is for the network based on the US 2002 IO table from Bureau of Economic Analysis (BEA, www.bea.gov).

features of IO networks without paying much attention to policy implications. Among the studies that do relate the structure features of IO networks to practical policy implications, their analyses are largely constrained by the limited number of network analysis tools applied, lacking a comprehensive understanding of the IO network. More specifically, most previous studies use national IO models, which undermine the policy relevance of their results in today's increasingly globalized economy.

Strongly rooted from the long legacy of IO analysis, this study examines the world economy as a network using the WIOD 2009 multi-regional input–output (MRIO) table by measuring the centrality of nodes and links, identifying communities in the world economy, and extracting backbone for the world economy. A comprehensive set of new approaches and indicators are proposed to better understand the structure of economies using IO networks. This study also tackled the intensive computational challenge posed by the much larger size of the IO network representing the world economy than those for national economies.

2. Methodology

We use commonly accepted symbols to describe our methods, where the economy has n sectors, x_{ij} measures the input from sector i to sector j, X_j represents the output of sector j, matrix A includes technical coefficients $a_{ij} = x_{ij}/X_j$, and L = (I-A)$^{-1}$ is the *Leontief inverse* matrix with I as the identity matrix [3]. An IO network can be constructed using the IO table of an economy. In particular, nodes are sectors and links are input–output transactions. The in-degree of a node k^{in} is the number of incoming links it has, while the out-degree k^{out} is the number of its outgoing links. Nodes can be weighted according to sector outputs, while links can be weighted using the value of input–output transactions they represent.

2.1. Network centrality

Particular nodes and links in a heterogeneous network play different roles, which can be characterized as their positions in the network. Nodes or links that are closer to the center of the network are intuitively more important than those are less close. In network analysis, betweenness and closeness are often used to measure the centrality of nodes or links. They are both based on the concept of shortest path which is the path connecting two particular nodes in the network with the least number of steps among all possible paths. The shortest path concept is suitable for most real-world networks studied by the literature, such as social networks, power grid, and transportation networks, in which the main functionality of networks is to efficiently mobilize information or objects. Thus, shorter paths are desired outcomes. For IO networks, however, the functionality is to maximize the value created by economic activities. In other words, we are interested in the circulation of monetary flows in the IO networks instead of quickly moving goods or exchanging services. Therefore, we compute betweenness and closeness differently in IO network analysis using the *Strongest Path* (SP) concept rooted from the structural path analysis (SPA) in IO modeling.

Centrality of a sector (node) in an IO network measures how "central" this sector is in the economy. Specifically, a sector that is closer to the center of the economy (higher centrality) by betweenness is more important in the sense that it facilitates more supply chains to make the economy thrive. A sector with a higher downstream or upstream closeness is more important to the economy because its products are supplies of more downstream sectors or its production generates more demand for upstream sectors, respectively.

2.1.1. Strongest path (SP)

SPA is used in IO analysis to measure contributions of particular supply chain path in the economy. In particular, to supply the production of sector j, there are multiple paths originating from all other sectors in the economy. Contributions of particular paths to the unitary output of sector j can be measured using the Taylor expansion of the *Leontief inverse* matrix [3].

Taking the contribution of sector i to sector j, among all possible paths, there is one particular path that contributes the most to the unitary output of sector j. We define this particular path from sector i to sector j as the *Strongest Path* (SP).[1] An SP represents the most important path of all possible paths that supply from a particular sector to another. For an IO network with n sectors, the number of SPs is n^2-n. The strength of a particular SP from sector i to sector j is measured as

$$q_{ij} = \prod a_{ik_1} a_{k_1k_2} \cdots a_{k_mj} \tag{1}$$

where the SP from sector i to j is identified as

$$P_{ij} = \{k_1, k_2, \cdots, k_m\} = i \to k_1 \to k_2 \to \cdots \to k_m \to j$$

where $i \neq k_1 \neq k_2 \neq \cdots \neq k_m \neq$h$j$. The number of direct paths that compose the SP is defined as the step of this particular SP (m+1 in this case).

Dijkstra's algorithm is used in network analysis for fast searching for shortest path between two nodes [49,51]. With slight modification, Dijkstra's algorithm can be used to search for SPs in IO networks. The modified Dijkstra-IO algorithm is described as pseudocode in Appendix 1.

2.1.2. SP betweenness and closeness

Betweenness in network analysis is defined as the number of shortest paths passing through a particular node or link. To take into account the strengths of SPs in IO networks, we define the SP betweenness of sector i as the weighted sum of strengths of all SPs in the IO network passing through it, not including SPs start or end at it:

$$b_i = \sum_{s=1, s\neq i}^{n} \sum_{t=1, t\neq i}^{n} X_t q_{st} \tag{2}$$

where $i \subset P_{st}$. Similarly, the SP betweenness for a particular link $i \to j$ is

$$b_{i\to j} = \sum_{s=1}^{n} \sum_{t=1}^{n} X_t q_{st} \tag{3}$$

where $i \to j \subset P_{st}$

Closeness in network analysis measures how far a particular node is to all other nodes based on their shortest paths. In IO networks, we define two SP-based closeness measures. In particular, downstream closeness[2] is the average value of all SPs starting from a particular sector i:

$$c_i^D = \frac{1}{n-1} \sum_{j=1}^{n} X_j q_{ij} \tag{4}$$

Similarly, upstream closeness[3] is defined as the average value of all SPs ending at a particular sector j:

$$c_j^U = \frac{1}{n-1} X_j \sum_{i=1}^{n} q_{ij} \tag{5}$$

[1] SP was previously termed as *Strongest inter-Sectoral Connection* or SSC in Xu et al. [10] M. Xu, B.R. Allenby, J.C. Crittenden, Interconnectedness and resilience of the U.S. Economy, Adv. Complex. Syst., 14 (2011) 649–672. For simplicity and easy interpretation, we re-termed it as SP in this paper.

[2] Previously defined as supply closeness in Xu et al. [10] ibid., re-termed for better interpretation.

[3] Previously defined as demand closeness in Xu et al. [10] ibid., re-termed for better interpretation.

Table 1
Popular community detection methods.

Method	Brief description
Hierarchical clustering	Define a similarity measure for nodes, group nodes with highest similarity, and then recalculate similarity to group other nodes, until all nodes are located in a single group [53].
Minimum cut	Divide the network into predetermined number of parts by minimizing the number of links between communities[54].
Girvan–Newman algorithm	Identify links with the highest betweenness and remove until the network split into sub components[27].
Modularity maximization	Group nodes connected by links that appear to be the most abnormal comparing to a randomly generated network with the same number of nodes and node degree (weighted number of links of each node) distribution [54].

Downstream closeness measures how close a particular sector is to its downstream consumers, while upstream closeness shows how close a sector is to its upstream suppliers. Another way to interpret these two closeness measures is that downstream closeness reveals the importance of a sector in the IO network as a supplier and upstream closeness does the same but treating the sector as a consumer.

2.2. Community

The heterogeneity of a network indicates the existence of a community (or cluster) structure within it [52]. Communities are constituted by strongly connected nodes. Many algorithms are available from the literature to detect communities in a network. Some of the popular ones are summarized in Table 1.

When determining the most appropriate algorithm for community detection, computational demand is a key factor, especially for IO networks. For example, the worst-case computational time required for the Girvan–Newman algorithm in its simplest and fastest form is $O(l^2 n)$ on a network with n nodes and l links.[4] It works for small or sparse networks. However, IO networks are usually dense in terms of the number of links per node. For instance, the IO network based on WIOD 2009 has 1435 nodes and 392,958 links, approximately 274 links per node, while the density of the Facebook social network is about 22 links per node, the citation network among US patents is around 4.4, the page link network from Google is 5.8, and the road network of Taxes is 2.8 or so [55]. Therefore we choose the modularity maximization algorithm [54] as it reduces the computational time requirement for community detection to $O((l+n)n)$. This allows reasonable times to complete running the algorithm, although still computationally intensive.

Formally, the modularity is defined to be

$$Q = \sum_h (e_{hh} - r_h^2) \tag{6}$$

where e_{hh} is the fraction of links that are in community h weighted using link strengths, r_h is the fraction of all ends of links that are connected to nodes in community h also weighted using link strengths, and r_h^2 therefore measures the weighted fraction of links connecting nodes in community h if the network is connected at random. High value of Q indicates high level of modularity in the network, thus a good community division. The modularity maximization algorithm finds the optimal community division for a network to maximize Q. The pseudocode of the modularity maximization algorithm is shown in Appendix 2.

2.3. Backbone

The heterogeneity of a network warrants heterogeneous importance of nodes and links. This leads to the question of what is the fundamentally essential structure of the network, or backbone, that non-essential components all attach to. Identifying the backbone of a network is particularly interesting for IO networks which are dense networks with hundreds of links per node. By eliminating less important links, one can find the most essential backbone of the IO network without being overwhelmed with massive amount of non-essential links.

The backbone of an economy extracted from the corresponding IO network reveals the most essential linkages of all sectors of the economy. The stability of this backbone makes the entire economy resilient even if some non-essential linkages are weakened or broken.

Early popular approaches to extract network backbone include the minimum spanning tree method which removes links of a network to minimize the total link weights while keeping all nodes connected [56]. However, this approach favors acyclic networks which may underrate the importance of local cycles in the original network. Another approach is to simply filter out links with weights below a predefined threshold, which ignore the fact that low weighted links may be relatively important for particular nodes [57,58]. In this study, we apply a disparity filter method [59] to address these methodological challenges for backbone extraction for IO networks.

[4] $O(z)$ is time complexity that measures the amount of time required to run an algorithm as the function of the size of its inputs (z).

Table 2
Sample six-sector input–output table ($).

Sector	1	2	3	4	5	6	Final demand	Total output
1	90	500	1000	0	0	10	300	1900
2	750	0	500	0	0	0	200	1450
3	650	750	0	190	0	0	175	1765
4	0	0	100	500	900	750	100	2350
5	0	0	0	800	390	510	300	2000
6	0	0	100	700	610	400	175	1985
Value added	410	200	65	160	100	315		
Total input	1900	1450	1765	2350	2000	1985		

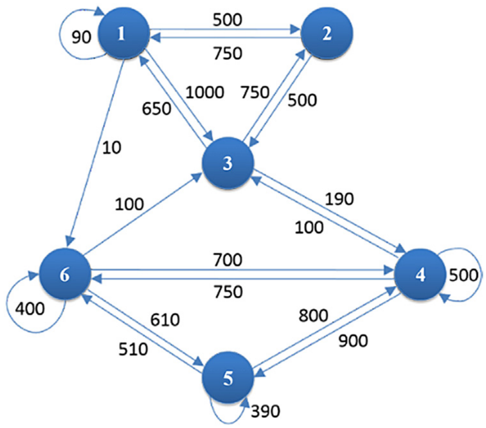


Fig. 2. Sample six-node input–output network showing link weights ($).

Disparity filter method compares the normalized weight of a link in the network with a null hypothesis that is the normalized weights of all links of a certain node are generated at random from a uniform distribution. If the null hypothesis is true, the probability of the normalized weight of a randomly generated incoming (outgoing) link larger than or equal to that of the observed link, known as the p value in statistical inference, can be calculated as

$$\alpha_{ij}^{in} = 1 - (k^{in} - 1) \int_{0}^{\hat{w}_{ij}^{in}} (1 - x)^{k^{in}-2} dx \tag{7}$$

$$\alpha_{ij}^{out} = 1 - (k^{out} - 1) \int_{0}^{\hat{w}_{ij}^{out}} (1 - x)^{k^{out}-2} dx \tag{8}$$

where k is the number of links the node has (degree) and $\hat{w}_{ij}$ is the normalized weight of the link i⟶j. Superscripts in and out indicate incoming and outgoing links, respectively. With a predefined significance level α, the null hypothesis is rejected when $\alpha_{ij} < \alpha$, implying that the link under consideration is significantly heterogeneous comparing its randomly generated counterpart. By imposing different significance levels, one can filter out non-essential links and keep the significantly heterogeneous links to extract the backbone of a network. The algorithm is described in Appendix 3 as pseudocode.

3. Sample network demonstration

We use a simple, hypothesized six-sector IO table to illustrate the IO network analysis described above (Table 2). A six-node network is constructed using the 6-by-6 IO matrix as the adjacency matrix (Fig. 2).

3.1. SP detection

By applying the Dijkstra-IO algorithm with technical coefficients, one can obtain all 30 (6×5) SPs. For example, the SP from sector 1 to sector 5 is 1⟶3⟶4⟶5, with three steps and the strength calculated by

$$q_{15} = \prod a_{13}a_{34}a_{45} = 0.567 \times 0.081 \times 0.450 = 0.021 \tag{9}$$

which means producing unitary output of sector 5 needs the largest amount of input from sector 1 through the link 1⟶3⟶4⟶5 than from any other possible links. Fig. 3 illustrates this by coloring the SP and showing technical coefficients associated with it.

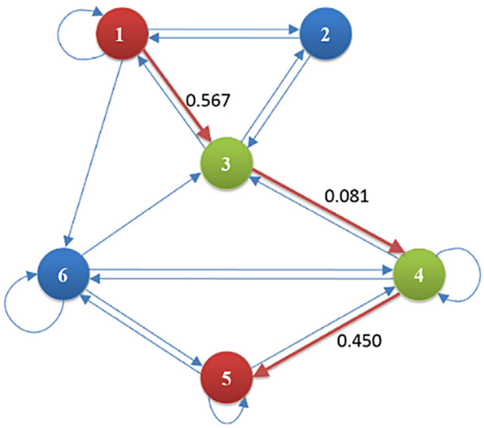


Fig. 3. SP from sector 1 to sector 5 showing technical coefficients associated with it.

Table 3
Node centrality of the sample input–output network ($).

Sector	SP betweenness	Downstream closeness	Upstream closeness
1	0	337.11	297.39
2	0	268.32	269.89
3	*460.49*	344.68	346.81
4	307.82	*365.86*	*370.29*
5	0	274.21	328.92
6	0	297.86	274.43

3.2. Centrality

Taking sector 4 as an example, there are nine SPs passing it: P_{15} (1→3→4→5), P_{16} (1→3→4→6), P_{25} (2→3→4→5), P_{26} (2→3→4→6), P_{35} (3→4→5), P_{36} (3→4→6), P_{51} (5→4→3→1), P_{52} (5→4→3→2), and P_{53} (5→4→3). Therefore, the SP betweenness of sector 4 is

$$b_4 = \sum_{s=1,S\neq 4}^{6} \sum_{t=1,t\neq 4}^{6} X_t q_{st} = 307.82(\$) \tag{10}$$

Considering the link 3→4, there are nine SPs passing through it: P_{14} (1→3→4), P_{15} (1→3→4→5), P_{16} (1→3→4→6), P_{24} (2→3→4), P_{25} (2→3→4→5), P_{26} (2→3→4→6), P_{34} (3→4), P_{35} (3→4→5), and P_{36} (3→4→6). The SP betweenness of link 3→4 is

$$b_{3\to 4} = \sum_{s=1}^{6} \sum_{t=1}^{6} X_t q_{st} = 598.25(\$) \tag{11}$$

Take sector 1 as an example for closenesses. There are five SPs starting from sector 1 (1→2, 1→3, 1→3→4, 1→3→4→5, and 1→3→4→6) and five SPs ending at it (2→1, 3→1, 4→2→1, 5→4→2→1, and 6→3→1). Therefore, the downstream closeness of sector 1 is

$$c_1^D = \frac{1}{5} \sum_{j=2}^{6} X_j q_{1j} = 337.11(\$) \tag{12}$$

and the upstream closeness is

$$c_1^U = \frac{1}{5} X_1 \sum_{i=2}^{6} q_{i1} = 297.39(\$) \tag{13}$$

Table 3 shows centralities of all sectors. In particular, sectors 3 and 4 are key sectors in this economy evidenced by their high SP betweenness values. Sector 3 seems to be more important than sector 4 justified by sector 3's higher SP betweenness. On the other hand, sector 4 is relatively more important than others as suppliers based on its downstream closeness, while also important as a consumer if considering upstream closeness. Table 4 shows SP betweenness of all links in the network. Links 1→s and 4→5 are more important than other links given their higher SP betweenness values.

Table 4
Link SP betweenness of the sample input–output network ($).

Sector	1	2	3	4	5	6
1	–	500.00	***1183.23***	–	–	–
2	750.00	–	591.62	–	–	–
3	736.19	849.45	–	598.25	–	–
4	–	–	240.37	–	***1034.61***	862.17
5	–	–	–	861.05	–	510.00
6	–	–	179.32	700.00	610.00	–

Table 5
Community detection processes using the modularity maximization algorithm.

Iteration	Q before iteration	△Q	Merged communities	Q after iteration	Sector					
					1	2	3	4	5	6
					Community					
0	0.0594	–	–	–	h_1	h_2	h_3	h_4	h_5	h_6
1st	0.0594	0.1309	h_1 and h_3	0.1903	h_1	h_2	h_1	h_3	h_4	h_5
2nd	0.1903	0.1229	h_1 and h_2	0.3132	h_1	h_1	h_1	h_2	h_3	h_4
3rd	0.3132	0.0624	h_2 and h_3	0.3756	h_1	h_1	h_1	h_2	h_2	h_3
4th	0.3756	0.00615	h_2 and h_3	0.4371	h_1	h_1	h_1	h_2	h_2	h_2
5th	0.4371	−0.4071	Stop	0.4371	h_1	h_1	h_1	h_2	h_2	h_2

Notes: The notation Q represents the modularity of the network. The modularity is a measure of the structure of networks. Networks with high modularity have dense connections between the nodes within same community but sparse connections between nodes in different communities.

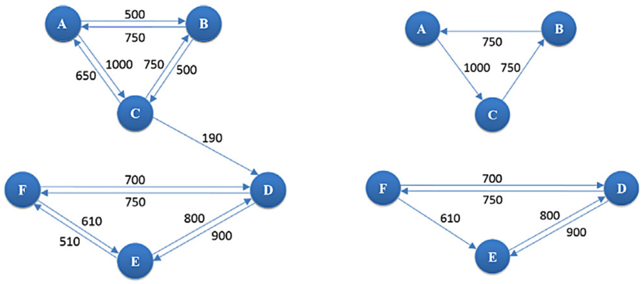


Fig. 4. Backbone of the sample input–output network for the significance level of 0.82 (left) and 0.30 (right).

3.3. Community

By applying the modularity maximization algorithm, two communities are detected from the sample IO network after four iterations (Table 5). In particular, sectors 1, 2 and 3 belong to one community while sectors 4, 5 and 6 form the other.

Notes: The notation Q represents the modularity of the network. The modularity is a measure of the structure of networks. Networks with high modularity have dense connections between the nodes within same community but sparse connections between nodes in different communities.

3.4. Backbone

The structure of the backbone of a network depends on the predefined significance level α. Fig. 4 shows two different backbone structures of the sample IO network for the significance level 0.82 and 0.30, respectively. Note that the backbone extracted under a lower significance level – more restrictive filtering – is part of the backbone extracted using a higher significance level meaning less restrictive filtering.

4. Network analysis for the world economy using WIOD 2009

We use the WIOD 2009 table to construct an IO network for the world economy. The WIOD 2009 table covers 41 economies (27 EU countries, 13 other major countries, and Rest of the World) each of which has 35 sectors. The resulted IO network has 1435 nodes and 392,958 links. It is a dense network with average in-degree/out-degree of approximately 274.

Table 6
Top ten sectors in the WIOD 2009 network by strongest path (SP) betweenness.

Rank	Country	Sector	SP Betweenness ($1,000,000)
1	RoW	Mining and Quarrying	87,580.29
2	US	Coke, Refined Petroleum and Nuclear Fuel	59,252.01
3	China	Basic Metals and Fabricated Metal	49,501.22
4	China	Coke, Refined Petroleum and Nuclear Fuel	43,994.56
5	China	Electrical and Optical Equipment	43,642.50
6	China	Mining and Quarrying	30,053.88
7	China	Chemicals and Chemical Products	28,652.30
8	Japan	Coke, Refined Petroleum and Nuclear Fuel	25,194.23
9	South Korea	Coke, Refined Petroleum and Nuclear Fuel	20,651.47
10	RoW	Basic Metals and Fabricated Metal	19,064.90

Table 7
Top ten sectors in the WIOD 2009 network by downstream closeness.

Rank	Country	Sector	Downstream closeness ($1,000,000)
1	US	Renting of Machinery and Equipment and Other Business Activities	1385.92
2	RoW	Mining and Quarrying	1225.43
3	US	Financial Intermediation	732.85
4	China	Basic Metals and Fabricated Metal	609.13
5	China	Chemicals and Chemical Products	406.84
6	China	Electrical and Optical Equipment	406.16
7	US	Real Estate Activities	397.54
8	China	Agriculture, Hunting, Forestry and Fishing	366.59
9	China	Mining and Quarrying	365.67
10	Japan	Renting of Machinery and Equipment and Other Business Activities	359.50

4.1. Centrality

Tables 6–8 show the top ten central sectors in the WIOD 2009 network measured by SP betweenness, downstream closeness, and upstream closeness. The full list of network centrality is included in Appendix 4 as an Excel file. The most central sector in the WIOD 2009 network is Mining and Quarrying in Rest of the World measured in SP betweenness. Nine out of the top ten sectors by SP betweenness are mining, energy, and basic materials, which reflects the importance of resources to the world economy. The fifth sector on the list is Electrical and Optical Equipment in China, which may be related to the booming of information and communications technology (ICT) and the fact that China is the main producer of ICT devices. In addition, five of the top ten sectors are from China, which indicates China's importance in the world economy as the center of transforming resources from all over the world into finished products to supply the globe.

The ranking of downstream closeness (Table 7) shows the world's top sectors as suppliers. It is interesting to note that Renting of Machine and Equipment and Other Business Activities in the US is the most important supplier in terms of downstream closeness among all sectors in the world. This may be mainly due to the "Other Business Activities" portion which includes a variety of service sectors in the US, showing the importance of the US service industry to the world economy. Other top suppliers on the list include resource extraction and material processing sectors in Rest of the World and China. Also included are Financial Intermediation and Real Estate Activities in the US, reflecting the importance of the US financial and real estate industries to the world. China's Electrical and Optical Equipment sector ranks the sixth on the list, probably also due to the booming of ICT consumption around the world. Japan's Renting of Machine and Equipment and Other Business Activities ranks the tenth, indicating Japan's service industry plays a key role in supporting the world economy.

Table 8 shows the top ten sectors ranked by upstream closeness, representing the world's most important sectors as consumers. Not surprisingly, the Public Administration and Defence & Compulsory Social Security sector in the US ranks the first, indicating its importance as a key consumer to drive the world economy. This is also evidenced by the fifth ranked Health and Social Work sector in the US. Construction sectors in China, Rest of the World, and the US rank the second, third, and eighth, respectively, which are also importance drivers to the world economy. This is particularly interesting given China's recent infrastructure and real estate development which led to massive booming in the construction industry. Intuition tells us that China's construction industry should be very important to the world economy, given the country's role in the world economy. Our results show that the IO network-based analysis can quantitatively identify this importance effectively. Also interested is the tenth ranked Food, Beverages and Tobacco in Rest of the World which indicates the importance of consumption to the world economy.

Note that many sectors appear in two or three lists, such as the Electrical and Optical Equipment sector in China. On the one hand, those sectors are probably more important than others ranked higher in only one of the lists, because they

Table 8
Top ten sectors in the WIOD 2009 network by upstream closeness.

Rank	Country	Sector	Upstream closeness ($1,000,000)
1	US	Public Admin and Defence; Compulsory Social Security	841.87
2	China	Construction	796.14
3	RoW	Construction	742.41
4	China	Electrical and Optical Equipment	587.03
5	US	Health and Social Work	448.01
6	China	Basic Metals and Fabricated Metal	445.04
7	US	Renting of Machinery and Equipment and Other Business Activities	444.24
8	US	Construction	430.33
9	US	Financial Intermediation	416.34
10	RoW	Food, Beverages and Tobacco	403.36

Table 9
Top ten communities by total output in the WIOD 2009 network.

Rank	Number of sectors	Community description
1	35	Manufacturing and services in the US, Canada, Mexico
2	33	Manufacturing and services in China
3	82	Manufacturing and services in RoW
4	33	Manufacturing and services in Japan
5	258	Manufacturing and services in Germany attached by Poland and Austria
6	34	Manufacturing and services in France
7	88	Manufacturing and services in UK attached by Ireland and Malta
8	39	Manufacturing and services in Italy attached by Malta's manufacturing
9	67	Manufacturing and services in Spain and Portugal
10	65	Manufacturing and services in Russia attached by Lithuania, Slovakia, and Sweden

function as both drivers and suppliers to the world economy. On the other hand, this also calls for better metrics to better quantify importance of sectors in the IO network, which represents an interesting future research avenue.

4.2. Community

Applying the modularity maximization algorithm results in 72 communities for the 1435-sector in WIOD 2009 network. The top ten communities by total output are shown in Table 9.

The world's largest economic cluster includes manufacturing and service sectors in North America, followed by the Chinese economy, RoW, and the Japanese economy. Each of the rest six clusters is dominated by one or two key countries in the European economy: Germany, France, the UK, Italy, Spain and Portugal, and Russia. In particular, the Germany-dominated cluster also includes sectors in Poland and Austria. Ireland and Malta are strongly connected with the UK, while some of Malta's manufacturing sectors are more closely attached with the Italian economy. Lastly, sectors in Lithuania, Slovakia, and Sweden are clustered with the Russian economy. The full list of clusters is included in Appendix 4.

4.3. Backbone

The structure of the backbone of a network depends on the predefined significance level α. Note that the backbone extracted under a lower significance level – more restrictive filtering – is part of the backbone extracted using a higher significance level meaning less restrictive filtering. Figs. 5 and 6 show the number and weights of links removed in the backbone of the WIOD 2009 network with increasing significance level. After the significance level increase to a certain threshold, the backbone of the network becomes stable as the number and weights of remaining links do not change with increasing significance level. For the WIOD 2009 network, the number of remaining links in the backbone becomes relatively stable when the significance level reaches to approximately 10^{-4}, while the weights of remaining links become stable when the significance level approaches 10^{-8}. Fig. 7 visualizes the backbone of the WIOD 2009 network with the significance level as 10^{-1} and 10^{-4}.

One of the benefits of extracting backbones for IO networks is that it allows faster computation of network analysis without losing important information. Table 10 shows top ten central sectors in terms of SP betweenness in the WIOD 2009 backbone network at the significance level of 10^{-4}. Comparing with Table 6 showing top sectors by SP betweenness in the original WIOD 2009 network, eight out of ten sectors appear in both lists.

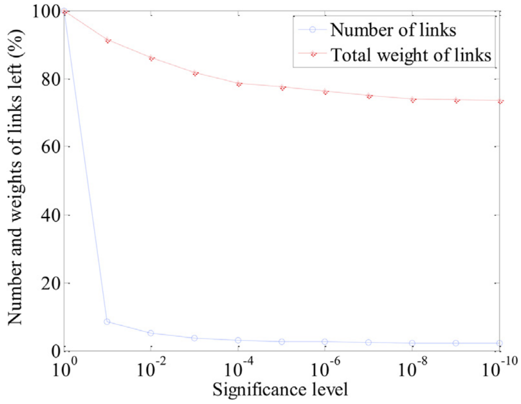


Fig. 5. Left in backbone extraction as the percentage of the total number and weights of links for the WIOD 2009 network at different significance levels.

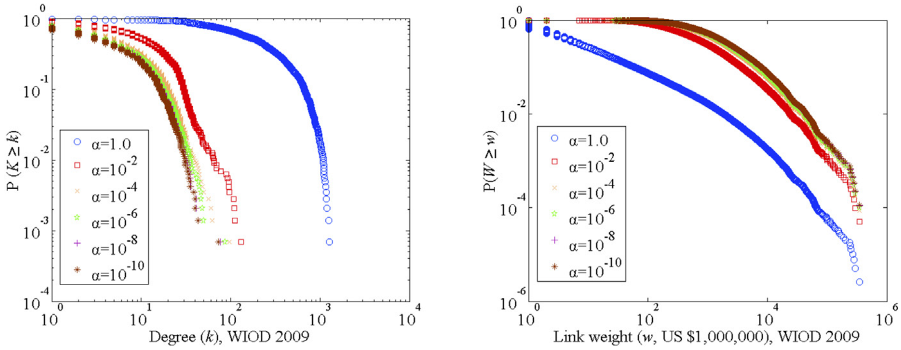


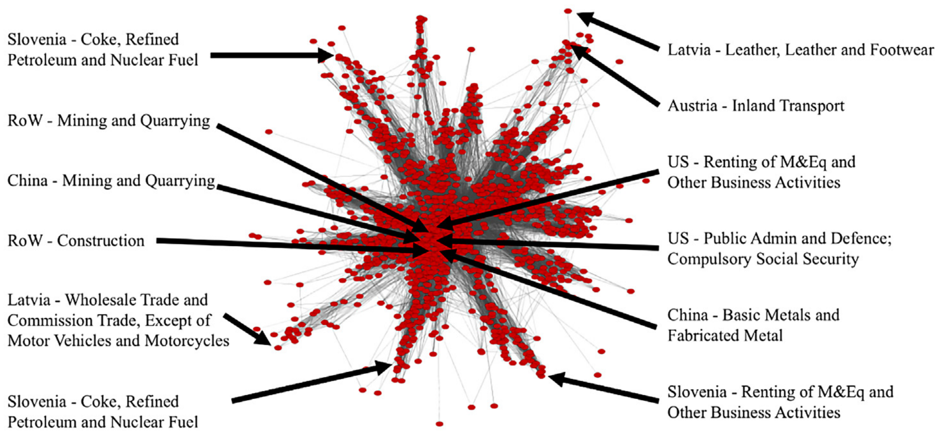
Fig. 6. Probability distribution of node degree (left) and link weight (right) in the backbone of the WIOD 2009 network at different significance levels.

Table 10
Top ten sectors by SP betweenness in the WIOD 2009 network backbone at the significance level of 10^{-4}.

Rank	Country	Sector	SP betweenness ($1,000,000)
1	RoW	Mining and Quarrying	408,741.37
2	China	Mining and Quarrying	163,889.91
3	US	Coke, Refined Petroleum and Nuclear Fuel	155,008.65
4	China	Basic Metals and Fabricated Metal	153,942.72
5	China	Electrical and Optical Equipment	138,717.38
6	China	Chemicals and Chemical Products	111,701.83
7	RoW	Basic Metals and Fabricated Metal	90,178.33
8	US	Renting of Machinery and Equipment and Other Business Activities	89,078.86
9	China	Coke, Refined Petroleum and Nuclear Fuel	86,309.21
10	US	Mining and Quarrying	80,290.01

5. Conclusions

We propose a network-based analysis using IO models to study structural characteristics of an economy. An IO network can be constructed based on an IO model in which nodes are sectors and directional, weighted links are IO transactions between sectors. We show that many IO networks at various scales possess heterogeneous properties that have been found in other real-world networks, which warrants the appropriateness of network analysis for IO-based economies. In this initial attempt, we use betweenness and closeness to measure the centrality of nodes and links in the IO network based on the strongest path concept which characterizes the most important supply chain between two sectors. The modularity maximization algorithm is used to detect communities in the IO network, identifying important economic clusters. Finally, a disparity filter method is used to extract backbones for the IO network at different significance level.



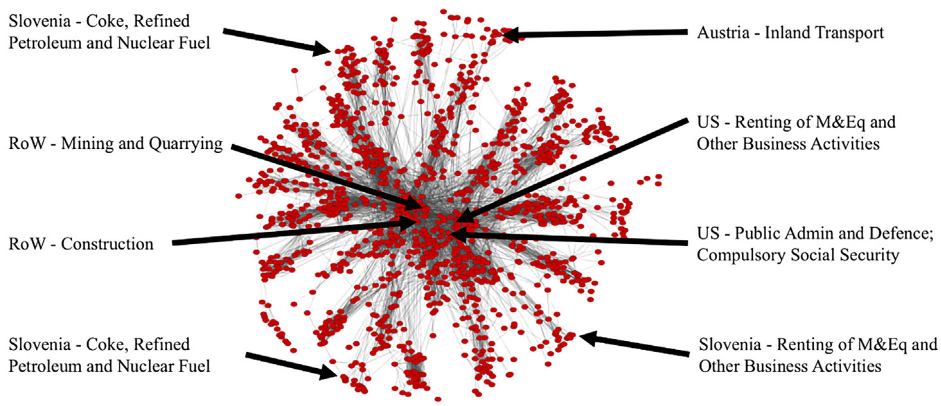


Fig. 7. Visualization of the WIOD 2009 network backbone at the significance level of 10^{-1} (upper) and 10^{-4} (lower). The RoW indicates the rest of the world.

We demonstrate the IO network analysis framework first using a hypothesized six-sector IO economy. The WIOD 2009 model is then used for a real-world case demonstration to examine the structure of the world economy. Our results show that the mining and quarrying industry of Rest of the World is the most central sector in the world economy. The US government services including public administration, defence, and compulsory social security are the most important driver for the world economy, while the US service industry is the most important supplier. Large communities detected include the North America economy, Chinese economy, Japanese economy, and economic clusters in Europe. Backbone of the world economy can be extracted using the disparity filter method to identify important links.

The results of our case study are intuitive. For example, energy-related sectors are generally located close to the center of the economy (high betweenness centrality), showing the importance of energy in facilitating supply chains. China's electrical and optical equipment sector has high betweenness centrality, corresponding to China's central role in the IT sector as the main producers of IT devices. Mining sectors are upstream to many other sectors, indicated by their high downstream closeness. Service sectors, especially those in large economies such as the US and China, are ultimate drivers of the economy by instigating upstream production with high upstream closeness. Communities detected by IO network analysis bring sectors from different countries together based on their fundamental economic relationship. These results provide a quantitative measure of many intuitive understandings of the economy and offer complementary insights on how the economy works at the global scale. Such insights are valuable for developing national economic policies with a structured, quantitative characterization of the structure of the economy.

The network representation of the IO model provides rich information for the structure of the IO economy and effective methods to characterize the structural properties. Additional insights are provided from the IO network analysis to better understand the structure of economies. This study lays the ground for future research and applications of IO network analysis. For example, a natural extension of this work is to examine the structural change of the global economy over time using time-series IO data.

Acknowledgments

This work was financially supported by the National Natural Science Foundation of China (71874014) and the Fundamental Research Funds for the Central Universities, China.

Appendix A. Supplementary data

Supplementary materials related to this article can be found online at https://doi.org/10.1016/j.physa.2019.121178.

References

[1] W. Leontief, Recent developments in the study of interindustrial relationships, Amer. Econ. Rev. 39 (1949) 211–225.
[2] W. Leontief, Input–output analysis as an alternative to conventional econometric models, Math. Soc. Sci. 25 (1993) 306–307.
[3] R.E. Miller, P.D. Blair, Input–Output Analysis: Foundations and Extensions, Cambridge University Press, New York, 2009.
[4] E. Dietzenbacher, B. Los, Structural decomposition techniques: Sense and sensitivity, Econ. Syst. Res. 10 (1998) 307–324.
[5] D. Guan, K. Hubacek, A new and integrated hydro-economic accounting and analytical framework for water resources: a case study for North China, J. Environ. Manag. 88 (2008) 1300–1313.
[6] M. Lenzen, Structural path analysis of ecosystem networks, Ecol. Model. 200 (2007) 334–342.
[7] S. Liang, T. Zhang, X. Jia, Clustering economic sectors in China on a life cycle basis to achieve environmental sustainability, Front. Env. Sci. Eng. 7 (2013) 97–108.
[8] G. Peters, Opportunities and challenges for environmental MRIO modeling: Illustrations with the GTAP database, in: 16th International Input-Output Conference of the IIOA, Istanbul, Turkey, 2007, pp. 64–69.
[9] S. Suh, Functions, commodities and environmental impacts in an ecological–economic model, Ecol. Econ. 48 (2004) 451–467.
[10] M. Xu, B.R. Allenby, J.C. Crittenden, Interconnectedness and resilience of the U.S. Economy, Adv. Complex. Syst. 14 (2011) 649–672.
[11] L. Chen, S. Liang, M. Liu, Y. Yi, Z. Mi, Y. Zhang, Y. Li, J. Qi, J. Meng, X. Tang, H. Zhang, Y. Tong, W. Zhang, X. Wang, J. Shu, Z. Yang, Trans-provincial health impacts of atmospheric mercury emissions in China, Nature Commun. 10 (2019) 1484.
[12] Z. Mi, J. Meng, F. Green, D.M. Coffman, D. Guan, China's exported carbon peak: Patterns, drivers, and implications, Geophys. Res. Lett. 45 (2018) 4309–4318.
[13] Z. Mi, J. Meng, D. Guan, Y. Shan, M. Song, Y.-M. Wei, Z. Liu, K. Hubacek, Chinese CO2 emission flows have reversed since the global financial crisis, Nature Commun. 8 (2017) 1712.
[14] J. Defourny, E. Thorbecke, Structural path analysis and multiplier decomposition within a social accounting matrix framework, Econ. J. 94 (1984) 111–136.
[15] E.J. Feser, E.M. Bergman, National industry cluster templates: A framework for applied regional cluster analysis, Reg. Stud. 34 (2000) 1–19.
[16] M. Sonis, G.J.D. Hewings, Economic complexity as network complication: Multiregional input–output structural path analysis, Ann. Reg. Sci. 32 (1998) 407–436.
[17] R. Wood, M. Lenzen, An application of a modified ecological footprint method and structural path analysis in a comparative institutional study, Local Environ. 8 (2003) 365–386.
[18] J. Camacho, R. Guimerà, L.A. Nunes-Amaral, Robust patterns in food web structure, in: J. Camacho, R. Guimerà, L.A. Nunes-Amaral (Eds.), Phys. Rev. Lett., 2002, p. 228102.
[19] H. Kitano, Systems biology: A brief overview, Science 295 (2002) 1662–1664.
[20] H. Jeong, B. Tombor, R. Albert, Z.N. Oltvai, A.L. Barabasi, The large-scale organization of metabolic networks, Nature 407 (2000) 651–654.
[21] A.L. Barabasi, R. Albert, Emergence of scaling in random networks, Science 286 (1999) 509–512.
[22] M.E.J. Newman, The structure and function of complex networks, Soc. Ind. Appl. Math. Rev. 45 (2003) 167–256.
[23] S.H. Strogatz, Exploring complex networks, Nature 410 (2001) 268–276.
[24] D.J. Watts, S.H. Strogatz, Collective dynamics of 'small-world' networks, Nature 393 (1998) 440–442.
[25] M.E.J. Newman, Networks: An Introduction, Oxford University Press, New York, 2010.
[26] G. Kossinets, D.J. Watts, Empirical analysis of an evolving social network, Science 311 (2006) 88–90.
[27] M. Girvan, M.E.J. Newman, Community structure in social and biological networks, Proc. Natl. Acad. Sci. USA 99 (2002) 7821–7826.
[28] A.E. Krause, K.A. Frank, D.M. Mason, R.E. Ulanowicz, W.W. Taylor, Compartments revealed in food-web structure, Nature 426 (2003) 282–285.
[29] F.Y. Wang, Parallel control and management for intelligent transportation systems: Concepts, architectures, and applications, IEEE Trans. Intell. Transp. Syst. 11 (2010) 630–638.
[30] A. Clauset, R. Guimerà, L.A. Nunes Amaral, Robust patterns in food web structure, Soc. Ind. Appl. Math. 51 (2002) 661–703.
[31] A. Klaus, S. Yu, D. Plenz, Statistical analyses support power law distributions found in neuronal avalanches, PLoS One 6 (2011) 1–12.
[32] M. Mitzenmacher, A brief history of generative models for power law and lognormal distributions, Internet Math. 1 (2004) 226–251.
[33] M.E.J. Newman, Power laws, pareto distributions and zipf's law, Contemp. Phys. 46 (2005) 323–351.
[34] J.M. Rueda-Cantuche, T.T. Raa, The choice of model in the construction of industry coefficients matrices, Econ. Syst. Res. 21 (2009) 363–376.
[35] M. Lenzen, D. Moran, K. Kanemoto, A. Geschke, Building EORA: A global multi-region input–output database at high country and sector resolution, Econ. Syst. Res. 25 (2013) 20–49.
[36] W. Liu, J. Chen, Z. Tang, H. Liu, D. Han, Y. Li, Multi-Regional Input–Output Table Establishment Theory and Application of China's 30 Provinces in 2007, China Statistics Press, Beijing, 2012.
[37] P. Slater, The network structure of the United States input–output table, Empir. Econ. 3 (1978) 49–70.
[38] F. BlöChl, T.J. Fabian, F.V. Redondo, E.O.N. Fisher, Vertex centralities in input–output networks reveal the structure of modern economies, in: F. BlöChl, T.J. Fabian, F.V. Redondo, E.O.N. Fisher (Eds.), Phys. Rev. E, 2011, p. 046127.
[39] S. Chen, B. Chen, Network environ perspective for urban metabolism and carbon emissions: A case study of vienna, Austria, Environ. Sci. Technol. 46 (2012) 4498–4506.
[40] S. Chen, B. Chen, Tracking inter-regional carbon flows: A hybrid network model, Environ. Sci. Technol. 50 (2016) 4731–4741.
[41] S. Chen, B. Xu, B. Chen, Unfolding the interplay between carbon flows and socioeconomic development in a city: What can network analysis offer? Appl. Energy 211 (2018) 403–412.

[42] S. Wang, Y. Liu, B. Chen, Multiregional input–output and ecological network analyses for regional energy–water nexus within China, Appl. Energy 227 (2018) 353–364.
[43] Y. Zhang, H. Zheng, Z. Yang, M. Su, G. Liu, Y. Li, Multi-regional input–output model and ecological network analysis for regional embodied energy accounting in China, Energy Policy 86 (2015) 651–663.
[44] J. McNerney, B.D. Fath, G. Silverberg, Network structure of inter-industry flows, Physica A 392 (2013) 6427–6441.
[45] S. Kagawa, S. Okamoto, S. Suh, Y. Kondo, K. Nansai, Finding environmentally important industry clusters: Multiway cut approach using nonnegative matrix factorization, Social Networks 35 (2013) 423–438.
[46] S. Kagawa, S. Suh, Y. Kondo, K. Nansai, Identifying environmentally important supply chain clusters in the automobile industry, Econ. Syst. Res. 25 (2013) 265–286.
[47] D. Acemoglu, V.M. Carvalho, A. Ozdaglar, A. Tahbaz-Salehi, The network origins of aggregate fluctuations, Econometrica 80 (2012) 1977–2016.
[48] D. Aobdia, J. Caskey, N.B. Ozel, Inter-industry network structure and the cross-predictability of earnings and stock returns, Rev. Account. Stud. 19 (2014) 1191–1224.
[49] E.W. Dijkstra, A Note on Two Problems in Connexion with Graphs, Springer-Verlag Inc., New York, 1959.
[50] W.Q. Duan, Modelling the evolution of national economies based on input–output networks, Comput. Econ. 39 (2012) 145–155.
[51] T.J. Misa, P.L. Frana, An interview with edsger W. Dijkstra, Commun. Assoc. Comput. Mach. 53 (2010) 41–47.
[52] M.E.J. Newman, Modularity and community structure in networks, Proc. Natl. Acad. Sci. USA 103 (2006) 8577–8582.
[53] D. Defays, An efficient algorithm for a complete link method, Comput. J. 20 (1977) 364–366.
[54] M.E.J. Newman, Fast algorithm for detecting community structure in networks, in: M.E.J. Newman (Ed.), Phys. Rev. E, 2004, p. 066133.
[55] SNAP, Stanford large network dataset collection, in: Stanford Network Analysis Project, 2013.
[56] M.L. Fredman, D.E. Willard, Trans-dichotomous algorithms for minimum spanning trees and shortest paths, J. Comput. System Sci. 48 (1994) 533–551.
[57] S. Allesina, A. Bodini, C. Bondavalli, Secondary extinctions in ecological networks: Bottlenecks unveiled, Ecol. Model. 194 (2006) 150–161.
[58] V.M. Eguíluz, D.R. Chialvo, G.A. Cecchi, M. Baliki, A.V. Apkarian, Scale-free brain functional networks, in: V.M. Eguíluz, D.R. Chialvo, G.A. Cecchi, M. Baliki, A.V. Apkarian (Eds.), Phys. Rev. Lett., 2003, p. 018102.
[59] M.A. Serrano, M. Boguñá, A. Vespignani, Extracting the multiscale backbone of complex weighted networks, Proc. Natl. Acad. Sci. USA 106 (2009) 6483–6488.