Structure of the Global Virtual Carbon Network

Revealing Important Sectors and Communities for Emission Reduction

Sai Liang, Yu Feng, and Ming Xu

Keywords:

carbon dioxide (CO_2)
complex networks
industrial ecology
input-output analysis (IOA)
supply chain
trade

Supporting information is available on the *JIE* Web site

Summary

Traditional production- and consumption-based accounting frameworks for carbon dioxide (CO_2) emissions focus on the two ends of supply chains, treating intermediate sectors as a "black box." Particular intermediate sectors can potentially be important for global mitigation of CO_2 emissions, through improving productivity to reduce inputs from upstream suppliers, thus emissions from upstream sectors, while still fulfilling downstream demands. Identifying those important intermediate sectors requires opening the black box and treating the economy as an integrated system. This study constructs a global virtual carbon network for 2009 and identifies key sectors for reducing global CO_2 emissions through improving productivity using network analysis techniques. We also identify 73 communities in the network in which sectors are more closely connected with one another than with sectors outside the community. Identifying communities helps in the understanding of potential impacts of sector-specific policy interventions through supply chains. The results offer additional insights that are not obviously visible in traditional input-output analysis.

Introduction

The specialization of production increases the dependence of sectors through the exchange of goods and services. These intersectoral linkages cause carbon leakages (i.e., carbon dioxide [CO_2] emissions embodied in exchanged goods and services) along global supply chains (Paltsev 2001). Although developed countries have made great efforts to achieve greenhouse gas (GHG) reduction goals of the Kyoto Protocol, and potentially Doha Amendment in the future, carbon leakages resulting from importing emission-intensive products from developing countries still drive the increase of global GHG emissions (Peters et al. 2011). To allocate global emissions to individual countries for developing international climate policies, both production- (assigning emissions to direct producers) and consumption-based (assigning emissions to final consumers) accounting frameworks are suggested (Davis and Caldeira 2010; Hertwich and Peters 2009).

Figure 1 shows a three-sector example to illustrate the production- and consumption-based accounting for CO_2 emissions. Suppose only sector 3 produces products that are used by final consumers. From a production-based perspective, sectors 1, 2, and 3 are responsible for CO_2 emissions they directly generate, c_1, 0, and c_3, respectively. From a consumption-based perspective, sector 3 is responsible for all emissions generated as a result of its production of final products ($c_1 + 0 + c_3$), where as sectors 1 and 2 do not have any consumption-based emissions. Consumption-based emissions are often estimated using environmentally extended input-output (I-O) models (Davis and Caldeira 2010; Peters 2008). These accounting results

Address correspondence to: Ming Xu, School of Natural Resources and Environment, University of Michigan, 440 Church Street, Ann Arbor, MI 48109-1041, USA. *Email:* mingxu@umich.edu

© 2015 by Yale University
DOI: 10.1111/jiec.12242

Editor managing review: Sybil Derrible

Volume 19, Number 2

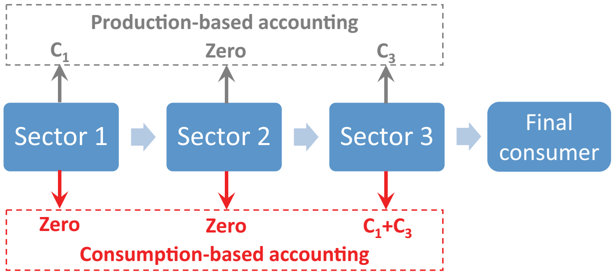

Figure 1 Three-sector example illustrating production- and consumption-based accounting of emissions.

can help develop policies focusing on either reducing carbon intensity in the production sectors or altering consumption patterns in the consumption sectors.

The production- and consumption-based emission accounting frameworks essentially identify either upstream sectors directly generating emissions (e.g., sector 1) or downstream sectors indirectly inducing the generation of emissions (e.g., sector 3) as important for climate policies. However, sectors in between often receive less attention. Potentially, those sectors along the supply chain can also be important for emission mitigation. For example, if sector 2 in figure 1 improves its production efficiency (i.e., using less inputs from sector 1 to produce the same amount of outputs for sector 3), it could reduce the production of sector 1, thus emissions from sector 1, while still providing the same amount of outputs for sector 3. Therefore, sector 2 can potentially contribute to the economy-wide emissions reduction by improving its production efficiency. Policies encouraging firms in sector 2 to improve production efficiency can be more acceptable and effective because they can also help reduce production cost by using less inputs from upstream suppliers. To identify those sectors that can help reduce the economy-wide emissions through improving production efficiency, one needs to examine specific supply chains instead of only focusing on aggregated production- or consumption-based emissions for individual sectors.

Structural path analysis (SPA), based on I-O models, is used to study supply chains in an economy at the sector level (Skelton et al. 2011). However, SPA focuses on measuring contributions of *separate* supply chain paths to *particular* sectors. It can identify important paths with respect to particular sectors (e.g., the path from sector 1 to 2 then to 3 is important to sector 3), but cannot identify key sectors important to the entire economy or the emissions of the entire economy (e.g., sector 2 to the economy-wide emissions). Identifying those sectors requires measuring their importance based on full intersectoral linkages, instead of individual supply chains.

In addition to the importance of individual sectors, it is also useful to identify clusters of sectors that can potentially reduce emissions collectively using the same or similar policy interventions (Kagawa et al. 2013). Sectors in the same cluster tend to have stronger relationships with one another than with sectors outside the cluster. The effects of policy interventions would spread faster within the same cluster. In this study, we aim to answer two research questions related to the structure of the global trade network and its environmental implications: (1) Which sectors are important to mitigating global CO_2 emissions through improving production efficiency? and (2) What are clusters of sectors that are important to mitigate global CO_2 emissions through improving production efficiency?

Modern network analysis (Barabási and Albert 1999; Newman 2003; Strogatz 2001; Watts and Strogatz 1998) offers an ideal framework for these questions. 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 systems (Newman 2010). In particular, a network consists of nodes (or vertexes) that are connected with one another through links (or edges). The structure of a network essentially reflects the specific way that nodes are connected by links. Network analysis provides a suite of techniques and metrics to comprehend the structure of a network and relate it to the functionality of the system represented by the network. In this study, we answer the above-mentioned research questions by examining the structural features of a global virtual carbon network built upon a global environmentally extended multiregional input-output (EE-MRIO) database using network analysis techniques.

Methods and Data

The Global Virtual Carbon Network

In an I-O model, an economy comprises sectors that are interdependent with each other through the exchange of goods and services. The I-O model tracks both direct and indirect supply-demand interdependencies among sectors within the entire economy. Multiregional input-output (MRIO) models are usually used to describe the global economy at the sector scale, such as Eora (Lenzen et al. 2013), the World Input-Output Database (WIOD) (Dietzenbacher et al. 2013), and the Global Trade and Analysis Project (Andrew and Peters 2013). An MRIO-based global economy can be regarded as a network in which nodes represent economic sectors, and links connecting nodes stand for the economic transactions between sectors (Nemeth and Smith 1985; Smith and White 1992). We can construct a global virtual carbon network based on the global MRIO model with a satellite account of CO_2 emissions. In this network, nodes are economic sectors, whereas links are virtual carbon flows among sectors (i.e., CO_2 emissions embodied in traded goods or services from one sector to another).

Based on the embodied emission concept (Subak 1995) and the EE-MRIO model, we represent the global virtual carbon Q using equation (1):

$$Q = \mathrm{diag}(E) * (I - A)^{-1} * \mathrm{diag}(y) \qquad (1)$$

where Q is a matrix with the element q_{ij} indicates the transfer of embodied CO_2 emissions from sector i to sector j to satisfy the final demand of products from sector j; vector E represents each sector's direct CO_2 emissions for one unit of its total output;

matrix A is the direct requirement coefficient matrix illustrating direct intersectoral relationships (Miller and Blair 2009); I is the identity matrix; matrix $(I-A)^{-1}$ is the Leontief inverse matrix characterizing both direct and indirect intersectoral relationships (Lenzen 2007; Miller and Blair 2009); vector y indicates the final demand of each sector's products; and the notation "diag" means diagonalizing the vector within the parenthesis.

The global virtual carbon network $C = (N, L)$ can be constructed based on Q. The set of nodes is expressed as $N = \{1, \ldots, n\}$, the set of directed links is expressed as $L = \{(i,j) \mid q_{ij} > 0\}$, and the weight of link (i,j) is q_{ij}.

This study is based on the WIOD (released in November 2013) (Dietzenbacher et al. 2013). The WIOD database covers 41 countries/regions (including 40 major countries/regions and the "rest of the world" [RoW]), 1,435 economic sectors (35 economic sectors per region), and a satellite account of sectoral CO_2 emissions. Treating sectoral CO_2 emissions as the satellite account of the MRIO table, we can construct the EE-MRIO model to describe relationships between the global economy and CO_2 emissions at the sector scale (Wiedmann et al. 2011). We construct the EE-MRIO model in this study using the WIOD 2009 data. Detailed descriptions of the EE-MRIO model and WIOD can be found in Dietzenbacher and colleagues (2013).

Centrality of the Global Virtual Carbon Network

A variety of real-world networks show strong heterogeneity (Clauset et al. 2009), which indicates that the network is not randomly connected and particular nodes or links play different roles in the network. In particular, different roles nodes or links playing can be topologically approximated by their positions in the network. *Betweenness* and *closeness* are two metrics 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 in the literature, such as social networks, power grid, and transportation networks. The main functionality of these networks is to efficiently mobilize information or objects. Taking the social network, for example, the benefit of forming a social network is to make the information exchange more efficient. In other words, a social network exists in order to transfer the information from one node to another at the highest possible efficiency. This efficiency can be considered as the ratio between the steps required to connect one node to another to the amount of the information transferred. The social network is unweighted, so the amount of the information transferred is always 1. Thus, the most efficient path connecting one node with another is the shortest path, and shortest paths are desired outcomes.

The specialization of production benefits countries by creating values (Hummels et al. 2001). Thus, for I-O networks, the benefit of forming a network is to create as much added value as possible. In other words, we are interested in the circulation of monetary flows in the I-O networks instead of quickly moving goods or exchanging services. In the IO-based global virtual carbon network, the cost of intersectoral connections is the amount of CO_2 emissions generated along the specific supply chain, whereas the benefit is the value added incurred along the supply chain. To identify the "hotspots" of CO_2 mitigation policies, we concern the most inefficient paths, which are paths causing the largest CO_2 emissions in the source industry i owing to 1 unit of value added created in the destination industry j. Therefore, we compute betweenness and closeness differently for the I-O-based global virtual carbon network using the *strongest path* (SP) concept. We recognize that there are many ways to describe the interactions among sectors. We choose the SP concept to demonstrate how modern network analysis tools can be applied to the analysis of IO-based networks, whereas using other metrics to describe the interactions among sectors remains as an interesting research avenue for the future.

Strongest Path

We define a path from sector i to sector j as a particular supply chain that starts from sector i and ends at sector j through a sequence of other sectors with any given sector appearing no more than once. Formally, a path from sector i to j is identified as $P_{ij}\{k_1, k_2, \ldots, k_m\} = i \rightarrow k_1 \rightarrow k_2 \rightarrow \ldots \rightarrow k_m \rightarrow j$, where $i \neq k_1 \neq k_2 \neq \ldots \neq k_m \neq j$ and $0 \leq m \leq (n-2)$. The number of direct connections that compose the path is defined as the step of this particular path $(m+1)$. For example, $P_{14}\{2\}$ represents the two-step path from sector 1 to sector 4 through sector 2, whereas $P_{14}\{\}$ stands for the direct path (one step) from sector 1 to sector 4 without passing through any other sectors. Each path characterizes a particular relationship between two sectors. In addition, paths containing circular loops are not considered in this study to avoid infinite searches for supply-chain paths. For instance, $P_{14}\{2,3,2\}$ is not a legitimate path because sector 2 appears more than once in the path. Allowing a sector appears no more than once in a path prevents circular supply chains from occurring (e.g., $1 \rightarrow 2 \rightarrow 3 \rightarrow 2 \rightarrow 4$ contains a circular supply chain $2 \rightarrow 3 \rightarrow 2$).

The number of all possible paths from sector i to sector j is given by equation(2):

$$\begin{aligned} p &= 1 + (n-2) + (n-2)(n-3) + \cdots + (n-2)! \\ &= \sum_{i=1}^{n-2} \frac{(n-2)!}{i!} \end{aligned} \tag{2}$$

Among all possible paths from sector i to sector j, there is one particular path that causes the largest CO_2 emissions in the source industry i owing to 1 unit of value added created in the destination industry j. We define this particular path as the SP from sector i to sector j. An SP represents the most inefficient path among all possible paths from a particular sector to another. For an IO-based network with n sectors, the number of SPs is $n(n-1) = n^2 - n$.

In particular, SPs are measured using the direct requirement coefficient matrix **A** (equation (3):

$$A = Z(\mathrm{diag}(x))^{-1} \tag{3}$$

where $n \times n$ matrix Z represents the intermediate intersectoral economic flows, and vector x indicates each sector's total output.

Equation (4) measures the strength of a particular path $P_{ij}\{k_1, k_2, \ldots, k_m\}$ in the global virtual carbon network:

$$q_{P_{ij}\{k_1,k_2,\cdots,k_m\}} = \frac{e_i a_{ik_1} a_{k_1k_2} \cdots a_{k_mj} x_j}{v_j} \tag{4}$$

where e_i is from vector E, $a_{ik_1}, a_{k_1k_2}, \ldots, a_{k_mj}$ are from matrix A, and v_j indicates value-added of sector j.

The SP from sector i to sector j in the global virtual carbon network is the path with the largest strength, expressed as SP_{ij}.

To compare roles of sectors playing in transmitting CO_2 emissions and economic values, this study also examines the centrality of the global economic network. The strength of a particular path $P^e_{ij}\{k_1, k_2, \ldots, k_m\}$ in the global economic network is measured by equation (5):

$$q_{P^e_{ij}\{k_1,k_2,\cdots,k_m\}} = \frac{a_{ik_1} a_{k_1k_2} \cdots a_{k_mj} x_j}{v_j} \tag{5}$$

Similarly, the SP from sector i to sector j in the global economic network is the path with the largest strength, expressed as SP^e_{ij}.

It is worth noting that the SP concept in this study is different from the concept of average propagation length (APL) (Dietzenbacher et al. 2005) in I-O analysis (IOA). The APL packs all intermediate paths between two nodes together and hence cannot support path-based centrality analysis. Dijkstra's algorithm is widely used for fast searching for shortest paths in a network (Dijkstra 1959; Thomas and Philip 2010). With slight modification, Dijkstra's algorithm can be used to search for SPs in an I-O-based global virtual carbon network and global economic network. The modified Dijkstra/I-O algorithm is described as the pseudocode in supporting information S1 on the Journal's website.

Strongest Path Betweenness and Closeness

Betweenness, in traditional 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 I-O-based networks, we define the SP betweenness of sector i as the weighted sum of the strengths of all SPs in the I-O-based network passing through it, excluding SPs' start or end at it. For sector i, its SP betweenness b_i is described by equation (6):

$$b_i = \sum_{s=1,s\neq i}^{n} \sum_{t=1,t\neq i}^{n} (v_t SP_{st}) \tag{6}$$

where $i \subset P_{st}$.

SP betweenness reveals the importance of a sector in the I-O-based network as a center transmitting or facilitating the creation of economic values or environmental impacts. A sector with high SP betweenness may not be large in terms of conventional metrics measuring the importance of sectors, such as direct energy use or emissions, total outputs, final demands, or value added. Therefore, SP betweenness allows for identifying those sectors that are important in transmitting or facilitating the creation of economic values or environmental impacts, but not obviously visible using traditional metrics. Identifying these sectors can potentially lead to the identification of important sectors for effective policy interventions.

Another commonly used metric for network centrality is closeness, which measures how far a particular node is to all other nodes based on their shortest paths. In I-O-based networks, we define two SP-based closeness measures. In particular, SP downstream closeness is the average value of all SPs starting from a particular sector i (equation (7):

$$c_i^D = \frac{1}{n-1} \sum_{j=1,j\neq i}^{n} (v_j SP_{ij}) \tag{7}$$

Similarly, SP upstream closeness is defined as the average value of all SPs ending at a particular sector j (equation (8):

$$c_j^U = \frac{1}{n-1} \sum_{i=1,i\neq j}^{n} (v_j SP_{ij}) \tag{8}$$

SP downstream closeness measures how close a particular sector is to its downstream consumers, whereas SP upstream closeness shows how close a sector is to its upstream suppliers. Another way to interpret these two closeness measures is that SP downstream closeness reveals the importance of a sector in the I-O-based network as a supplier (providing more inputs for other sectors), and SP upstream closeness identifies important sectors as consumers (consuming more outputs of other sectors).

Two concepts in IOA are similar to the two closeness metrics: production-based accounting and consumption-based accounting (Peters 2008). SP downstream closeness is similar to production-based accounting, which allocates emissions to producers. SP upstream closeness is similar to consumption-based accounting, which allocates emissions to consumers. This study calculates production- and consumption-based CO_2 emissions to compare with results based on SP down- and upstream closeness.

Community Detection for the Global Virtual Carbon Network

A lot of real-world networks exhibit robust community structure (Newman 2006). In particular, communities are clusters of nodes that are closely connected with one another, but loosely connected with others. Changes in individual nodes of a network are more likely to affect those nodes that are in the same community. For the global virtual carbon network, a community represents a cluster of sectors that affect one another's CO_2 emissions more critically than sectors outside the community. In other words, changes in a particular sector are likely to affect more on emissions generated by other sectors in the same community. Those sectors in the same community may not necessarily belong to the same country or same geographic region.

Table 1 Popular community detection methods

Methods	*Brief descriptions*
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 cluster (Defays 1977).
Minimum cut	Divide the network into predetermined number of parts by minimizing the number of links between communities (Newman 2004a).
Girvan-Newman method	Identify links with the highest betweenness and remove until the network is split into subcomponents (Girvan and Newman 2002).
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 (Newman 2004b)

Therefore, identifying communities in the global virtual carbon network is important for understanding potential impacts of policy interventions at the global scale.

There are I-O-based cluster analyses from the literature, such as the restricted maximization method (Hoen 2002) and minimal flow analysis (Titze et al. 2011). These methods are based on filter factors specified exogenously and ignore link weights. Network-based approaches do not need an arbitrary filter factor and take into account link weights to identify communities. In particular, many approaches in network analysis are available to detect communities in a network. Some of the popular ones are summarized in table 1.

When determining the most appropriate approach for community detection, computational demand is a key factor, especially for I-O-based networks that are densely connected. For example, the worst-case computational time required for the Girvan-Newman method (Girvan and Newman 2002) in its simplest and fastest form is $O(l^2n)$ on a network with n nodes and l links. $O(z)$ is time complexity that measures the amount of time required to run an algorithm as the function of the size of its inputs (z). This approach works for small or sparse networks. However, I-O-based networks are usually dense in terms of the average number of links per node. For instance, the I-O-based network based on WIOD 2009 has 1,435 nodes and 1,777,454 links, approximately 1,239 links per node, whereas the density of the Facebook social network is approximately 22 links per node, the citation network among U.S. patents is approximately 4.4, the page link network from Google is 5.8, and the road network of Texas is 2.8 or so (SNAP 2013). Therefore, we choose the modularity maximization approach (Newman 2004b) because it reduces the computational time requirement for community detection to $O((l+n)n)$. This allows reasonable time to complete running the algorithm.

Formally, modularity is defined by equation (9):

$$M = \sum_h (e_{hh} - r_h^2) \tag{9}$$

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 M indicates high level of modularity in the network, thus a good community division. The modularity maximization approach finds the optimal community division for a network by maximizing M. The pseudocode for modularity maximization shown in the Supporting Information on the Web is designed according to Newman's fast algorithm for detecting community structure (Newman 2004b).

Note that there is no universally applicable approach for network community detection. The modularity maximization approach used in this study may have a number of potential pitfalls (Good et al. 2010). Random null model in the modularity maximization approach is assumed to be a fully connected network (Good et al. 2010; Fortunato 2010). This assumption is generally consistent with the characteristic of I-O networks, in which exchanges of goods or services exist almost between any two sectors. In addition, the modularity maximization approach may fail to identify small communities (Fortunato and Barthélemy 2007), which is less relevant for this study given that we are mainly interested in large communities in the global virtual carbon network. Developing better techniques for community detection for I-O-based networks represents an interesting future research avenue.

Results

The global virtual carbon network based on 2009 WIOD data has 1,435 nodes and 1,852,228 directed, weighted links. The global economic network also has 1,435 nodes, but slightly less number of links (1,777,454). Links in the global virtual carbon network reflect both direct and indirect connections between any two nodes through global supply chains. Although there is no direct economic connection between two particular nodes (node i and node j), node i can transfer embodied CO_2 emissions to node j through indirect connections (e.g., a path starting from node i, ending at node j, and passing through other nodes). Subsequently, the global virtual carbon network has more links than the global economic network.

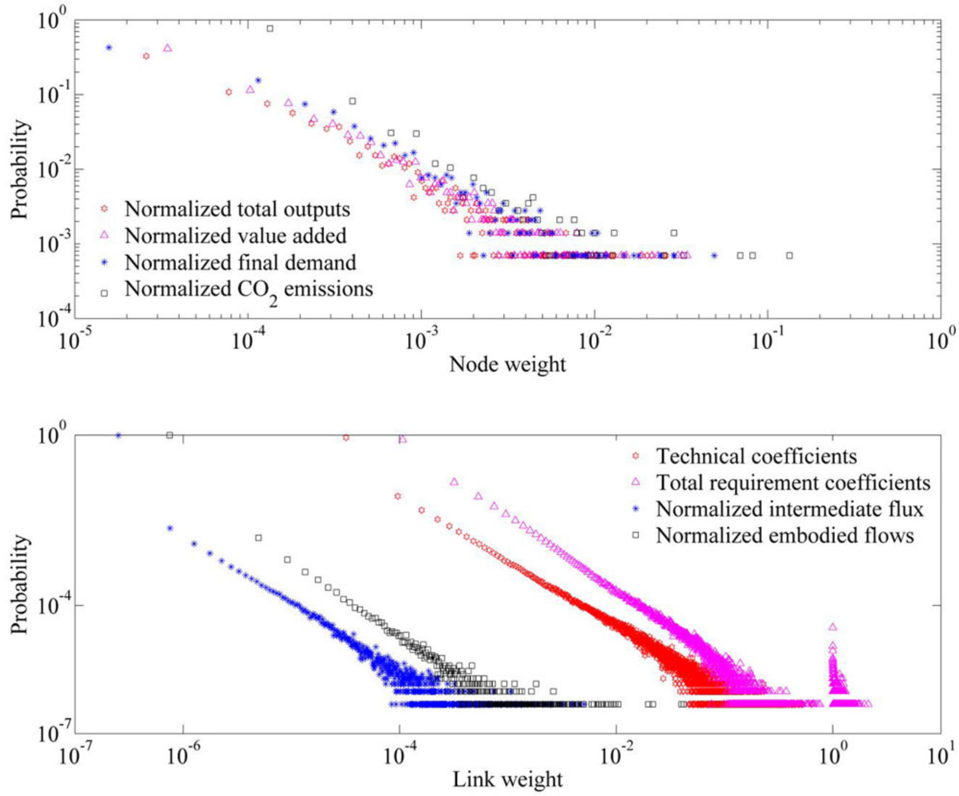

Figure 2 Probability distribution of the weight of nodes and links in the global virtual carbon network. Intermediate flux and sectoral total outputs are normalized by global total economic outputs. Sectoral final demands and value added are normalized by global gross domestic product (GDP). Embodied carbon flux and sectoral CO_2 emissions are normalized by total global CO_2 emissions. The probability distribution of these indicators follows the Power Function which is $f(x) = ax^b + c$. Values for parameters *a*, *b*, and *c* are listed in table S1-1 in the supporting information on the Web. CO_2 = carbon dioxide.

Heterogeneity of the Global Virtual Carbon Network

The structure of 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 (Clauset et al. 2009; Klaus et al. 2011; Mitzenmacher 2004; Newman 2005). Note that nodes and links could also be heterogeneous themselves in many ways (e.g., size and strength). Such heterogeneity indicates that components of a system play different roles in the system, which thus implies that the structure of the system is critical to the functionality of the system.

We plot the probability distribution of node and link weights measured by a variety of metrics for the global virtual carbon network (figure 2). In particular, the probability distributions tell the likelihood that a node or link with particular weights appears in the network. We use economic attributions to measure nodes and links, include sectoral total output, value added, and final demand for nodes, and intermediate flux, technical coefficients, and total requirement coefficients for links. Also, sectoral CO_2 emissions and embodied CO_2 flux are used to weight nodes and links, respectively. Figure 2 shows that small nodes and weak links are more likely to appear in the network than large nodes and strong links. This indicates the heterogeneity of the global virtual carbon network, implying that nodes and links play different roles in the network.

In common sense, larger nodes and stronger links are more important. However, network analysis offers more insights through centrality metrics, betweenness and closeness in particular, beyond the simple comparison of the weight of nodes and links.

Centrality of Nodes and Links in the Global Virtual Carbon Network

We calculate SP betweenness, SP downstream closeness, and SP upstream closeness for each sector in the global virtual carbon network and global economic network to measure centrality of nodes. The full results are shown in the Supporting Information on the Web. Figure 3 shows the probability distribution of the centrality of sectors in the global virtual carbon network and global economic network. A disproportionally small portion of sectors are in the center of the networks (i.e., large centrality values but low probability), indicating that a few sectors dominate the transmission of embodied carbon flows and economic flows.

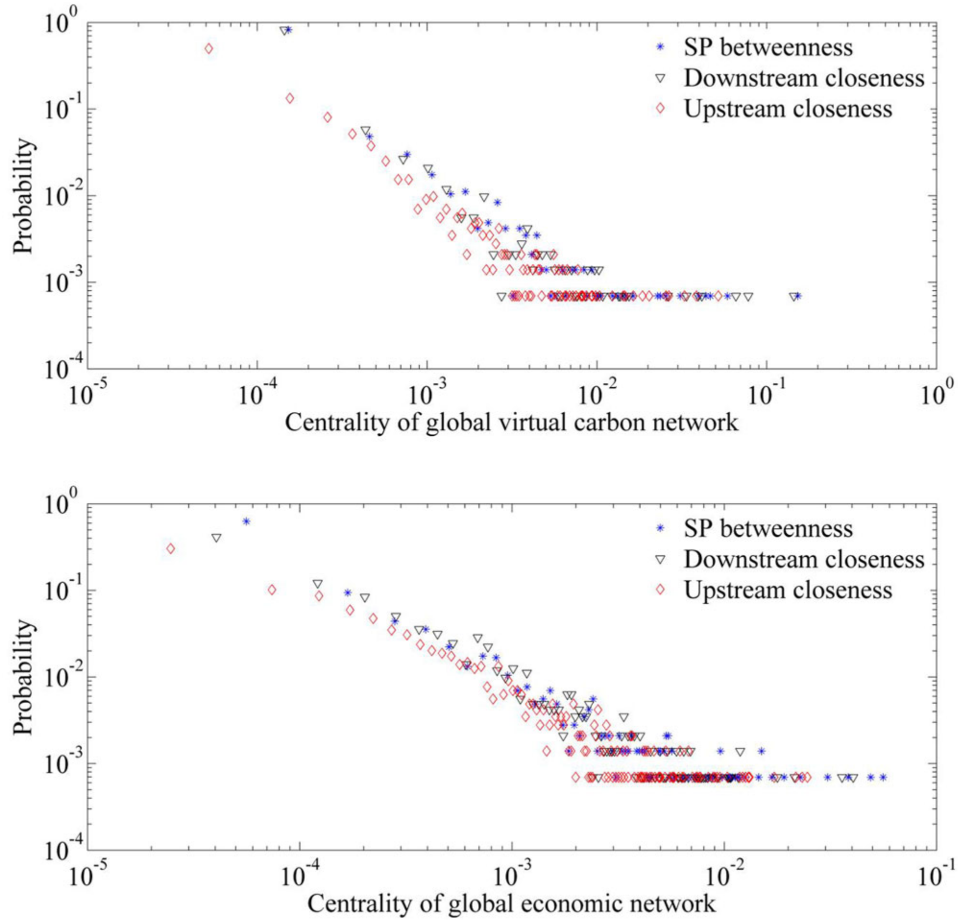

Figure 3 Probability distribution of the centrality of sectors in the global virtual carbon network and global economic network. Centrality values are normalized by the sum of the particular metric. The probability distribution of these metrics follows the power function, which is $f(x) = ax^b + c$. Values for parameters *a*, *b*, and *c* are listed in table S1-1 in the supporting information on the Web. SP = strongest path.

The SP closeness results validate the rationality of SP-based centrality analysis. Table S1-2 in the supporting information on the Web shows the top 15 sectors by SP downstream closeness in the global virtual carbon network and global economic network. SP downstream closeness measures the importance of a sector as a supplier in the networks, or to say, major emitters in the global virtual carbon network and key producers in the global economic network. Therefore, it is expected that the ranking by SP downstream closeness should be similar to that by production-based accounting, as evidenced by table S1-2 and correlation coefficients in table S1-5 in the supporting information on the Web. Table S1-3 in the supporting information on the Web shows the top 15 sectors by SP upstream closeness in the global virtual carbon network and global economic network. SP upstream closeness describes the importance of a sector as a consumer driving CO_2 emissions and value-added creations. Therefore, it is expected that the ranking by SP upstream closeness should be similar to that by consumption-based accounting, as generally evidenced by table S1-3 in the supporting information on the Web and correlation coefficients in table S1-5 in the supporting information on the Web.

Figure 4 shows the ranking of the top 15 sectors with respect to SP betweenness in the global virtual carbon network and global economic network (a full visualization of the ranking of all sectors can be found in the Supporting Information on the Web). These sectors are the most importance ones in terms of transmitting and facilitating the generation of CO_2 emissions and economic values. Thirteen of the top 15 sectors are related to resource extraction and processing, including mining, energy production, and basic materials production, reflecting the importance of resources for the functioning of the global economy and the generation of global CO_2 emissions. Those central sectors of the global virtual carbon network mainly locate in China, Russia, the United States, India, Japan, Taiwan, and the RoW. The most important center of the global virtual carbon network is mining and quarrying in the RoW. The most important center of the global economic network is coke, refined petroleum, and nuclear fuel in the United States, indicating the importance of the U.S. energy sector in the world economy. In particular, electrical and optical equipment in China is the sixth-most central sector of the global virtual carbon network and also the fourth-most central sector of the global economic network,

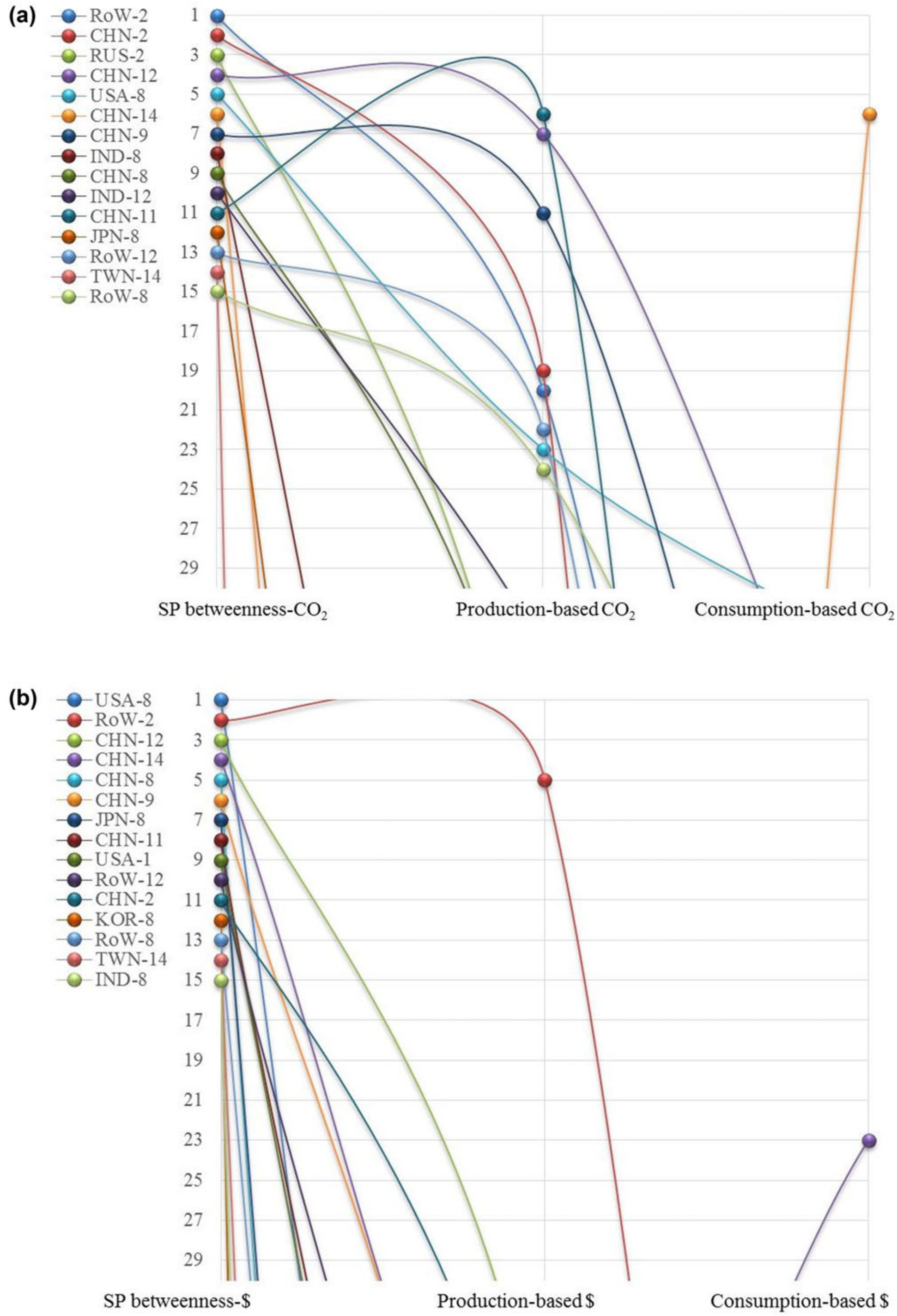

Figure 4 Ranking of sectors in (a) the global virtual carbon network and (b) the global economic network, showing only the top 15 sectors by SP betweenness. Sectors are described by a three-letter code for the country and an index for the sector. Full names of country/region abbreviations and sector indices are listed in supporting information S2 on the Web. Absolute values this figure is based on are listed in Table S1-4 in supporting information S1 on the Web. SP = strongest path; CO_2 = carbon dioxide.

which may be owing to China's role as the main producer of information and communication equipment. In addition, six of the top 15 central sectors in the two networks are from China, indicating China's role in the world economy as the center of transforming global resources into finished products for global consumption, which also leads to great global CO_2 emissions. Moreover, mining and quarrying in Russia and basic metals and fabricated metal in India rank top in the global virtual carbon network, but not in the top 15 for the global economic network, indicating that these sectors facilitate a large amount of CO_2 emissions, but do not create equivalently large economic values. This implies that policy interventions to these sectors and their supply chains can potentially make significant impacts to reduce the amount of CO_2 emissions generated in the corresponding supply chains and create more-economic values.

We also compare the ranking of sectors by SP betweenness, production-based CO_2 emissions, and consumption-based CO_2 emissions for the global virtual carbon network. Top central sectors by SP betweenness are generally not top ones in terms of either production- or consumption-based CO_2 emissions. Similarly, top central sectors by SP betweenness in the global economic network are generally not top sectors measured by either production- or consumption-based value added. For example, electric and optical equipment in Taiwan ranks among the top 15 in both the global virtual carbon network and the global economic network, whereas it only ranks 712th in production-based CO_2 emissions, 320th in consumption-based CO_2 emissions, 373rd in production-based value added, and 448th in consumption-based value added, of a total of 1,435 sectors. Low correlation coefficients between SP betweenness of sectors and I-O-based metrics validate such findings (table S1-5 in the supporting information on the Web). These results indicate the additional insights that SP betweenness can bring to the traditional IOA.

Tables 2 and 3 show the top 15 links (direct exchange of goods or services from one sector to another) with respect to SP betweenness in the global virtual carbon network and global economic network. These links are the most important intersectoral connections that help transmit and facilitate the generation of CO_2 emissions and economic values. These top 15 links within the global virtual carbon network are fossil fuel (FF) mining to processing, fuel consumption by transport, usage of nonmetallic minerals by construction, and raw materials supply to the manufacturing of equipment and machinery. The most important link as a transmission center within the global virtual carbon network is from mining and quarrying to coke, refined petroleum and nuclear fuel within China, indicating the importance of China's FF processing in global CO_2 emissions. On the other hand, these top 15 links within the global economic network are fuel consumption by public administration and inland transport, the usage of nonmetallic mineral and metals by construction, metals usage by equipment and machinery, FF mining to processing, and agricultural product supply to foods production. The most important link is from coke, refined petroleum, and nuclear fuel to public administration and defense and compulsory social security within the United States.

In particular, 8 of the top 15 links within the global virtual carbon network and 9 of the top 15 links within the global economic network are related with China, indicating China's role as the "world factory" in the global economy.

We also find that top central links by SP betweenness within these two global networks are generally not top ones in terms of embodied CO_2 flows and direct economic flows (also validated by low correlation coefficients in table S1-6 in the supporting information on the Web), indicating that SP betweenness can bring additional insights to the traditional IOA.

Improving the productivity of sectors and links with high SP betweenness (e.g., sectors and links related with basic materials production in China) in the global virtual carbon network can help mitigate global carbon emissions.

Communities of the Global Virtual Carbon Network

To identify major sector clusters causing global CO_2 emissions, we divide the global virtual carbon network into 73 communities by the modularity maximization approach, with the modularity value as 0.76. Each community represents a cluster of sectors that are strongly connected with one another by carbon leakages, but weakly connected with other sectors in the network. Table 4 lists the top 25 communities measured by the sum of CO_2 emissions generated by all sectors in each community.

The largest community consists of mining, manufacturing, construction, and services sectors in China, with several textiles, leather, and equipment sectors in some European countries. This community generates 5.88 billion tonnes (t) of CO_2, representing 23.6% of the global total. The second-largest community is the RoW without agriculture, attached by transportation sectors in major European countries and metals production in Turkey, with 4.30 billion t of CO_2 emissions (17.3% of the global total). The U.S. mining, manufacturing, construction, and services sectors without water and air transport constitute the third-largest community with 3.86 billion t of CO_2 emissions (15.5% of the global total). The descriptions for all 73 communities are shown in supporting information S2 on the Web.

Whereas some communities are consistent with the geographic boundaries of countries (e.g., Australia and Brazil are basically communities themselves), most large communities are not limited by geographic constraints. For example, agriculture, foods, and leather sectors in the United States are closer to Canada (community 8) than other U.S. sectors (community 3) in the global virtual carbon network.

Discussion and Conclusions

This study applies a variety of network analysis tools to uncover the structural features of the global virtual carbon network, with comparison with the global economic network. Probability distributions of the weight of nodes and links measured by various metrics indicate that the two networks are highly heterogeneous, in the way that a few nodes or links are disproportionally larger or stronger than the majority of nodes or links.

We identify hotspots in the global virtual carbon network using three centrality metrics: betweenness, downstream closeness, and upstream closes, all based on the concept of SP. The results identify those sectors acting as important

Table 2 Top 15 links by SP betweenness in the global virtual carbon network

	Global virtual carbon network					
	Origin sector		*Destination sector*			
Rank	*Country/ region*	*Sector name*	*Country/ region*	*Sector name*	*SP betweenness ($kt\ CO_2$)*	*Rank in embodied CO_2 flows (out of 2 million)*
1	China	Mining and quarrying	China	Coke, refined petroleum, and nuclear fuel	60,384	2,256
2	RoW	Mining and quarrying	USA	Coke, refined petroleum, and nuclear fuel	25,314	577
3	RoW	Mining and quarrying	RoW	Coke, refined petroleum, and nuclear fuel	20,677	733
4	India	Coke, refined petroleum, and nuclear fuel	India	Inland transport	17,325	335
5	China	Other nonmetallic mineral	China	Construction	14,082	3
6	China	Basic metals and fabricated metal	China	Electrical and optical equipment	12,577	35
7	RoW	Mining and quarrying	China	Coke, refined petroleum, and nuclear fuel	11,676	9,924
8	RoW	Mining and quarrying	Japan	Coke, refined petroleum, and nuclear fuel	11,665	1,276
9	USA	Coke, refined petroleum, and nuclear fuel	USA	Public administration and defense and compulsory social security	11,067	114
10	RoW	Mining and quarrying	South Korea	Coke, refined Petroleum, and nuclear fuel	10,992	2,615
11	China	Coke, refined petroleum, and nuclear fuel	China	Water transport	8,379	2,284
12	China	Basic metals and fabricated metal	China	Machinery	8,198	47
13	Russia	Inland transport	RoW	Coke, refined petroleum, and nuclear fuel	8,120	678
14	Taiwan	Electrical and optical equipment	China	Electrical and optical equipment	7,244	8,052
15	China	Electrical and optical equipment	RoW	Electrical and optical equipment	6,468	5,202

Note: RoW = rest of the world; SP = strongest path; kt CO_2 = kilotonnes of carbon dioxide.

Table 3 Top 15 links by SP betweenness in the global economic network

	Global economic network					
	Origin sector		*Destination sector*			
Rank	*Country/ region*	*Sector name*	*Country/ region*	*Sector name*	*SP betweenness (million US$)*	*Rank in economic flows (out of 2 million)*
1	USA	Coke, refined petroleum, and nuclear fuel	USA	Public administration and defense and compulsory social security	36,625	96
2	China	Other nonmetallic mineral	China	Construction	36,491	11
3	China	Basic metals and fabricated metal	China	Electrical and optical equipment	34,927	17
4	China	Mining and quarrying	China	Coke, refined petroleum, and nuclear fuel	29,712	45
5	USA	Agriculture, hunting, forestry, and fishing	USA	Food, beverages, and tobacco	29,571	22
6	China	Basic metals and fabricated metal	China	Machinery	22,255	23
7	USA	Mining and quarrying	USA	Coke, refined petroleum, and nuclear fuel	21,716	34
8	China	Basic metals and fabricated metal	China	Construction	17,834	12
9	RoW	Mining and quarrying	USA	Coke, refined petroleum, and nuclear fuel	17,734	47
10	China	Agriculture, hunting, forestry, and fishing	China	Food, beverages, and tobacco	17,126	9
11	China	Chemicals and chemical products	China	Rubber and plastics	16,997	46
12	USA	Coke, refined petroleum, and nuclear fuel	USA	Inland transport	16,823	248
13	India	Coke, refined petroleum, and nuclear fuel	India	Inland transport	16,703	166
14	China	Coke, refined petroleum, and nuclear fuel	China	Inland transport	16,244	299
15	Taiwan	Electrical and optical equipment	China	Electrical and optical equipment	15,620	164

Note: RoW = rest of the world; SP = strongest path; US$ = U.S. dollars.

Table 4 Top 25 communities in the global virtual carbon network

Rank	*CO_2 emissions (kt CO_2)*	*Community*
1	5,877,800	Mining, manufacturing, construction, and services in China, attached by textiles, leather, and equipment in several European countries
2	4,303,800	Rest of the world, except for agriculture, attached by transportation activities in major European countries and metals production in Turkey
3	3,864,500	Mining, manufacturing, construction, and services in the USA, except for water and air transport
4	1,501,200	India, attached by mining in Belgium and chemicals production in Malta
5	1,448,800	Russia and Slovak Republic, attached by inland transport in Lithuania and fossil fuel processing in Slovenia and Sweden
6	1,279,900	Major European countries (i.e., Austria, Belgium, Czech Republic, Germany, Denmark, France, Hungary, Luxembourg, Netherlands, and Slovenia)
7	974,920	Japan, attached by water transport in China
8	550,250	Canada, attached by agriculture, foods, and leather in the USA
9	504,520	South Korea, except for water transport
10	441,480	United Kingdom and Ireland, attached by financial intermediation in Luxembourg, papermaking in Malta, and fossil fuel processing in Netherlands
11	369,330	Italy, Malta, and Slovenia, attached by air transport in France and Hungary
12	364,320	Australia
13	348,210	Mexico, except for electrical and optical equipment
14	331,190	Indonesia, except for sale, maintenance, and repair of motor vehicles and motorcycles as well as retail sale of fuel
15	282,040	Spain and Portugal
16	274,460	Poland, except for textiles and leather production
17	251,470	Taiwan, except for textiles and water transport
18	251,290	Brazil
19	242,220	Agriculture-related activities in China
20	237,350	Water transport in rest of the world
21	230,500	Turkey, except for basic metals and fabricated metal
22	229,750	Textiles and water transport in Taiwan and agriculture-related activities in rest of the world
23	216,420	Bulgaria, Cyprus, Greece, and Romania, attached by coking in Malta
24	157,510	Denmark, Estonia, Finland, Lithuania, Latvia, and Sweden
25	155,830	Air transport in USA

Note: kt CO_2 = kilotonnes of carbon dioxide.

transmission centers, suppliers, and consumers in the global virtual carbon network, providing additional insights that traditional IOA cannot offer. Those sectors should be the focus of sustainable supply-chain management for global CO_2 mitigation. For example, sectors with high SP betweenness are important transmission hubs and, potentially, have significant leverage power to reduce global CO_2 emissions through improving their production efficiency. Moreover, sectors with high centrality in the global economic network, but low centrality in the global virtual carbon network, generally deserve promotion given that, relatively, they are associated with more-economic values, but less CO_2 emissions. We also identify communities in the global virtual carbon network. Sectors in the same community are more likely affected by one another in terms of CO_2 emissions. Identifying communities helps understand potential impacts of policy interventions targeting at particular sectors, given that the impact of policies can be more effectively transferred through the close connections among sectors within the same community. Moreover, community detection in this study reveals strongly connected sectors that are from different countries, offering additional insights regarding the structure of the world economy that traditional country-based IOA cannot easily provide.

In addition to the network analysis metrics used in this study, there are other techniques and tools in network analysis that can potentially be applied in IOA. Introducing more network analysis tools to uncover the structural features of the global virtual carbon network in future work can provide additional information for international climate-change policy making. In addition to the global virtual carbon network studies in this work, one can construct similar networks for major resources (e.g., metals, water, and materials) and other emissions of global importance. Moreover, network analysis is not yet an area where one can take a method “off the shelf.” Much theoretical and methodological explorations are needed for future studies.

Acknowledgments

The material is based upon work supported by the U.S. National Science Foundation (NSF) under Grant No. 1438197.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF. Sai Liang thanks the support of the Dow Sustainability Fellows Program.

References

Andrew, R. M. and G. P. Peters. 2013. A multi-region input-output table based on the global trade analysis project database (GTAP-MRIO). *Economic Systems Research* 25(1): 99–121.

Barabási, A.-L. and R. Albert. 1999. Emergence of scaling in random networks. *Science* 286(5439): 509–512.

Clauset, A., C. Shalizi, and M. Newman. 2009. Power-law distributions in empirical data. *Society for Industrial and Applied Mathematics (SIAM) Review* 51(4): 661–703.

Davis, S. J. and K. Caldeira. 2010. Consumption-based accounting of CO_2 emissions. *Proceedings of the National Academy of Sciences of the United States of America* 107(12): 5687–5692.

Defays, D. 1977. An efficient algorithm for a complete link method. *The Computer Journal* 20(4): 364–366.

Dietzenbacher, E., I. R. Luna, and N. S. Bosma. 2005. Using average propagation lengths to identify production chains in the Andalusian economy. *Estudios de Economía Aplicada* 23(2): 405–422.

Dietzenbacher, E., B. Los, R. Stehrer, M. Timmer, and G. de Vries. 2013. The construction of world input-output tables in the WIOD project. *Economic Systems Research* 25(1): 71–98.

Dijkstra, E. W. 1959. A note on two problems in connexion with graphs. *Numerische Mathematik* 1(1): 269–271.

Fortunato, S. 2010. Community detection in graphs. *Physics Reports* 486(3–5): 75–174.

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

Girvan, M. and M. E. J. Newman. 2002. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences of the United States of America* 99(12): 7821–7826.

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

Hertwich, E. G. and G. P. Peters. 2009. Carbon footprint of nations: A global, trade-linked analysis. *Environmental Science & Technology* 43(16): 6414–6420.

Hoen, A. R. 2002. Identifying linkages with a cluster-based methodology. *Economic Systems Research* 14(2): 131–146.

Hummels, D., J. Ishii, and K.-M. Yi. 2001. The nature and growth of vertical specialization in world trade. *Journal of International Economics* 54(1): 75–96.

Kagawa, S., S. Okamoto, S. Suh, Y. Kondo, and K. Nansai. 2013. Finding environmentally important industry clusters: Multiway cut approach using nonnegative matrix factorization. *Social Networks* 35(3): 423–438.

Klaus, A., S. Yu, and D. Plenz. 2011. Statistical analyses support power law distributions found in neuronal avalanches. *PLoS ONE* 6(5): e19779.

Lenzen, M. 2007. Structural path analysis of ecosystem networks. *Ecological Modelling* 200(3–4): 334–342.

Lenzen, M., D. Moran, K. Kanemoto, and A. Geschke. 2013. Building Eora: A global multi-region input-output database at high country and sector resolution. *Economic Systems Research* 25(1): 20–49.

Miller, R. E. and P. D. Blair. 2009. *Input-output analysis: Foundations and extensions* (2nd edition). Cambridge, UK; New York: Cambridge University Press.

Mitzenmacher, M. 2004. A brief history of generative models for power law and lognormal distributions. *Internet Mathematics* 1(2): 226–251.

Nemeth, R. J. and D. A. Smith. 1985. International trade and world-system structure: A multiple network analysis. *Review (Fernand Braudel Center)* 8(4): 517–560.

Newman, M. 2003. The structure and function of complex networks. *Society for Industrial and Applied Mathematics (SIAM) Review* 45(2): 167–256.

Newman, M., ed. 2010. *Networks: An introduction*. New York: Oxford University Press.

Newman, M. E. J. 2004a. Detecting community structure in networks. *The European Physical Journal B–Condensed Matter and Complex Systems* 38(2): 321–330.

Newman, M. E. J. 2004b. Fast algorithm for detecting community structure in networks. *Physical Review E* 69(6): 066133.

Newman, M. E. J. 2005. Power laws, Pareto distributions and Zipf's law. *Contemporary Physics* 46(5): 323–351.

Newman, M. E. J. 2006. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences of the United States of America* 103(23): 8577–8582.

Paltsev, S. V. 2001. The Kyoto Protocol: Regional and sectoral contributions to the carbon leakage. *The Energy Journal* 22(4): 53–79.

Peters, G. P. 2008. From production-based to consumption-based national emission inventories. *Ecological Economics* 65(1): 13–23.

Peters, G. P., J. C. Minx, C. L. Weber, and O. Edenhofer. 2011. Growth in emission transfers via international trade from 1990 to 2008. *Proceedings of the National Academy of Sciences of the United States of America* 108(21): 8903–8908.

Skelton, A., D. Guan, G. P. Peters, and D. Crawford-Brown. 2011. Mapping flows of embodied emissions in the global production system. *Environmental Science & Technology* 45(24): 10516–10523.

Smith, D. A. and D. R. White. 1992. Structure and dynamics of the global economy: Network analysis of international trade 1965–1980. *Social Forces* 70(4): 857–893.

SNAP (Stanford Network Analysis Platform). 2013. Stanford large network dataset collection. http://snap.stanford.edu/data. Accessed.

Strogatz, S. H. 2001. Exploring complex networks. *Nature* 410(6825): 268–276.

Subak, S. 1995. Methane embodied in the international trade of commodities: Implications for global emissions. *Global Environmental Change* 5(5): 433–446.

Thomas, J. M. and L. F. Philip. 2010. An interview with Edsger W. Dijkstra. *Communications of the ACM* 53(8): 41–47.

Titze, M., M. Brachert, and A. Kubis. 2011. The identification of regional industrial clusters using qualitative input-output analysis (QIOA). *Regional Studies* 45(1): 89–102.

Watts, D. J. and S. H. Strogatz. 1998. Collective dynamics of 'small-world' networks. *Nature* 393(6684): 440–442.

Wiedmann, T., H. C. Wilting, M. Lenzen, S. Lutter, and V. Palm. 2011. Quo Vadis MRIO? Methodological, data and institutional requirements for multi-region input-output analysis. *Ecological Economics* 70(11): 1937–1945.

About the Authors

Sai Liang is a Dow Sustainability Postdoctoral Fellow, **Yu Feng** was a master's student, and **Ming Xu** is an assistant professor, all in the School of Natural Resources and Environment at the University of Michigan, Ann Arbor, MI, USA. **Ming Xu** is also an assistant professor in the Department of Civil and Environmental Engineering at the University of Michigan.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web site:

Supporting Information S1: This supporting information S1 contains the parameters of the power function for probability distribution in figures 2 and 3 of the main article, the pseudocode for the Dijkstra/I-O algorithm for detecting SPs within the global virtual carbon network, the pseudocode for the modularity maximization approach for community detection, and results for centrality analysis of the global virtual carbon network and the global economic network.

Supporting Information S2: This supporting information S2 provides definitions of all country abbreviations and sector numbers, full centrality results for sectors within the global virtual carbon network and the global economic network, full results for SP betweenness of links within the global virtual carbon network and the global economic network, the correspondence between sectors and communities (with descriptions of communities), and the ranking of all country sectors according to various indicators.