Using Network Science to Quantify Economic Disruptions in Regional Input-Output Networks

Emily P. Harvey and Dion R. J. O'Neale

Abstract Input-output (IO) tables provide a standardised way of looking at inter-connections between all industries in an economy, and are often used to estimate the impact of disruptions or shocks on economies. IO tables can be thought of as networks—with the nodes being different industries and the edges being the flows between them. We develop a network-based analysis to consider a multi-regional IO network at regional and sub-regional level within a country. We calculate both linear matrix-based IO measures ('multipliers') and new network theory-based measures, and contrast these measures with the results of a disruption model applied to the same IO network. We find that path-based measures (betweenness and closeness) identify the same priority industries as the simulated disruption modelling, while eigenvector-type centrality measures give results comparable to traditional IO multipliers, which are dominated by overall industry strength.

1 Introduction

Economic disruptions such as those due to natural hazards have a large impact on local and global economies. There is evidence that the flow-on impacts of disruptions will have an increasing impact as the world becomes more globalised and inter-connected [1, 4]. In order to build resilience and prioritise investment to mitigate impacts, it is crucial to identify key industry sectors and regions that play

E. P. Harvey (✉)
M.E Research, Takapuna, Auckland, New Zealand

Te Pūnaha Matatini, Centre of Research Excellence, Auckland, New Zealand
e-mail: emily@me.co.nz

D. R. J. O'Neale
Physics Department, University of Auckland, Auckland, New Zealand

Te Pūnaha Matatini, Centre of Research Excellence, Auckland, New Zealand
e-mail: d.oneale@auckland.ac.nz

a role in amplifying (or dampening) the flow-on impacts of disruptions or shocks. When considering disruption impacts, economists and disruption planners are faced with the need to evaluate a number of different measures of economic impacts that are not necessarily comparable and which almost certainly can not all be minimised simultaneously.

Internationally, studies of flow-on impacts of disruptions on economic systems have most commonly been based on the data from input-output (IO) tables, which are readily available, at least at a national, and often also at regional levels. Many years of research has gone into using IO tables in economic impact analysis [14] and a common approach is to use 'multipliers', based on linear algebra matrix formulations, to estimate the indirect impacts of a change in demand (or supply) for an industry. In response to natural hazard events, the most popular approach has been inoperability input-output models, which are based on standard 'IO multiplier analysis' with minor modifications [6].

Recently IO tables have been thought of as networks, with the nodes being the different industries and the edges being the flows between the industries. This has enabled network science techniques to be used to attempt to identify crucial nodes (industry sectors) within an economy and other industry structures. Existing work in this field began with calculating the properties of IO networks [2, 3, 13], and is now beginning to investigate the propagation of shocks on the networks [4, 11, 13, 16].

In this work we seek to consider the connections between industries as a network, to determine whether the structure of the network can provide useful extra information for quantifying the importance of industries or regions in propagating disruptions. We consider a multi-regional IO network at local (Territorial Authority) level within the Waikato Region in New Zealand, and calculate both linear matrix-based IO measures (e.g. 'multipliers') and network theory-based measures at this higher spatial resolution. We compare these network-based measure with results from a disruption model applied to the same IO data, which gives us further information about disruption impacts.

Research to date has considered global IO networks, looking at flows within and between each country, or looking at a single country of interest. However, when considering more fine-grained economic data regions are often heterogeneous and impacts can be highly local. By comparing and contrasting the analysis at both Regional and Local spatial resolutions, we are also able to investigate the impact of spatial resolution on the results obtained.

2   Background and Data

2.1   Level of Spatial and Industry Aggregation

In this work, our starting data is a Multi-regional Input-Output table (MRIO) which partitions the 10 Territorial Authorities (TAs) in the Waikato Region into separate subregions and breaks the rest of New Zealand into 'North of the Waikato Region'

(Auckland and Northland) and 'South of the Waikato Region' (all other Regions). This gives us 12 different spatial regions which span a large range of sizes, both geographically and economically. The number of industry sectors is aggregated to 106, which is the maximum allowed from the reference data used to construct the IO tables.

2.2 Economic Network Setup

IO networks are weighted, directed networks, where the weighting indicates the size ($) of inter-industry-location flows, and the direction depends on which industry-location the flow is from and to. In this work we have 1272 industry-location nodes (12 regions (locations) and 106 (industry) sectors) in the network. In addition to the flows between industry sectors, the IO tables we use include inputs from four Value Added categories (labour and capital inputs, taxes/subsidies on products and production) and Imports, and outputs to three Final Demand categories (household and government consumption, capital formation) and Exports. This adds another three nodes to each of the 12 'regions' for Final Demand, four nodes for Value Added at whole country level, and a node each for Imports and Exports.

It is a requirement of IO tables that flows in and out of an industry-location must balance, so this places restrictions on the network, specifically the row and column sums of the weights in the adjacency matrix must much once value added and final demand components are included. Additionally, because the nodes are the grouping of all industries of the same 'sector' in the TA, self-links are possible and will typically account for a significant fraction of monetary flows. There are some other features of IO networks that it is worth noting; one of which is that they are very dense (nearly-complete) with most industries having connections with most other industries, though not all of these flows (link weights) will be significant. These features mean that a lot of the standard approximations and simplifications for weighted, directed network analysis are often not able to be applied [12].

3 Method

3.1 Network Analysis

In network science *centrality* broadly refers to the 'importance' or 'influence' of a node in the network. Centrality measures can range from local node properties (e.g. node strength); to more extensive properties that consider the properties of those nodes connected to a node of interest (e.g. PageRank centrality); to measures that account for the structure or topology of the network (e.g. betweenness centrality). In this work we analyse the network using a range of different centrality measures, including those that have been identified as potentially important in

economic networks. Where possible we consider the network as a weighted, directed network, with self-loops, but not all algorithms allow for this.

The centrality measures used here were chosen to cover a range different types of measures while keeping the range of measures manageable. Specifically, we consider the following (from the igraph package [5]):

Node Strength An equivalent of node degree that accounts for the differing edge weights in a network. The strength of a node is simply the sum of the weights of edges connected to that node.

PageRank A popular variant of eigenvector centrality. Eigenvector-based measures consider not only the strength of a node but also how well connected a node is to other nodes with high node strength.

Kleinburg Centrality A generalisation of eigenvector centrality. Nodes are imbued with two attributes: *Authority*—how much information/influence is held by the node; and *Hubness*—how well a node connects to nodes with high authority. If A is the adjacency matrix of a network, the hub score of nodes is calculated as the principle eigenvector of AA^T.

Authority Related to hub centrality, the authority centrality of vertices is calculated as the principle eigenvector of $A^T A$.

Diversity The scaled Shannon entropy of the edge weights of a node. Given the context of IO networks it is worth noting that the diversity measure used here is a version of the species diversity measure commonly used in ecology to quantify the diversity of a habitat [9], not the measure of diversity sometimes used in economic geography and popularised by Hausmann and Hidalgo in [8] which is simply the node degree of a binary matrix that measures whether a region is strongly associated with particular products or exports.

Betweenness The number of weighted shortest paths that pass through a node, given all possible paths between pairs of nodes in the network.

Closeness The number of steps required to reach every other node in a network from a given node following edge-weighted paths.

Having calculated this selection of centrality measures, we use Kendall's τ [10] to calculate the correlation between the importance rankings of the industries between any two centrality measures.

3.2 Multiplier Analysis

IO Multipliers In calculating the economic effects of changes in an economy (positive or negative shocks), 'IO multipliers' are the most commonly used approach. When the change is considered on the demand side the 'Leontief inverse' is used, which propagated the change to all downstream industries. When the change

is considered on the supply side the 'Ghosh inverse' is used which effectively propagates the change to all upstream industries. See [14] for a full description.

The main technique used for quantifying economic impacts of a disaster (for example a natural hazard such as an earthquake or volcanic eruption) is known as Inoperability Input-Output Model (IIM) or the Dynamic IIM as a time varying extension [6]. In this model, the inoperability of industries is assumed to follow a smooth logistic curve from the disaster induced loss of productive capacity back to full capacity over a specified recovery period. The direct loss of production due to industry-location inoperability is calculated and used to modify the final demand by the same amount; that is, if production halved from $20,000 to $10,000 then the final demand vector for that industry-location in that region would be reduced by $10,000. Then the flow on impacts from this reduced demand would be calculated using the Leontief inverse [14]. This continues through time, until full operability is restored.

In this work we calculate both Type I (industry to industry spending only) multipliers and Type II (including household spending and labour income) multipliers, for the whole multi-regional IO table, following [14].

Disruption Multipliers There are many issues with this IIM approach [15], including that it can lead to double counting and not only inaccurate quantitative results, but more importantly it can lead to inaccurate rankings for prioritisation of industries. Harvey et al. [7] have instead developed a dynamic model that propagates short-term (days to weeks) disruptions through the multi-regional IO network. Using this model 'disruption multipliers' can be calculated by disrupting one industry-location at a time in each region and working out the ratio of direct effects to flow-on (indirect) effects throughout the whole of New Zealand.

3.3 Comparing Network Centralities with IO and Disruption Multipliers

We use Kendall's τ [10] to calculate the correlation between industry-location rankings based on the multiplier measures compared to the centrality measures.

3.4 Comparing Spatial Aggregation

In parallel with this, we also construct a network at the level of the Waikato Region (not separated at TA level) and the same 'North of the Waikato Region' and 'South of the Waikato Region' regions (3 network regions) to investigate the impact of spatial aggregation of IO tables on the economic multipliers and the network properties.

4   Results

4.1   *Network Centrality Measures and Their Correlations*

Figure 1 shows a heatmap of the Kendall correlation coefficients [10] between all the centrality measures considered here. Before comparing the centrality measure rankings, we first remove the industries from the rest of New Zealand (North and South of the Waikato Region), as these have much larger inputs/outputs than those broken down by TA within the Waikato Region and risk dominating the results. Furthermore, we are focused here on identifying important industry-location pairs within the Waikato Region.

   We find that overall the different eigenvector-based centrality measures are highly correlated, in particular the Kleinburg Authority, Kleinburg Hub, and PageRank, and that these are strongly driven by the node strength (total inputs/outputs) of the industry-location pairs. The different path-based measures are also highly correlated, for example, the closeness and betweenness measures. These path-based measures highlight different industry-location pairs to those identified by the eigenvector-based methods, as shown by the high level of anti-correlation (red) between these types of measures. More importantly, the path-based measures identify industry-location pairs that would not be immediately revealed by eigenvector-based methods, or by linear economic multiplier measures that are strongly linked to their size (strength) in the local economy. We elaborate on this in the next section.

Fig. 1 Centrality measure correlations (τ values) for industries in the ten TAs in the Waikato Region
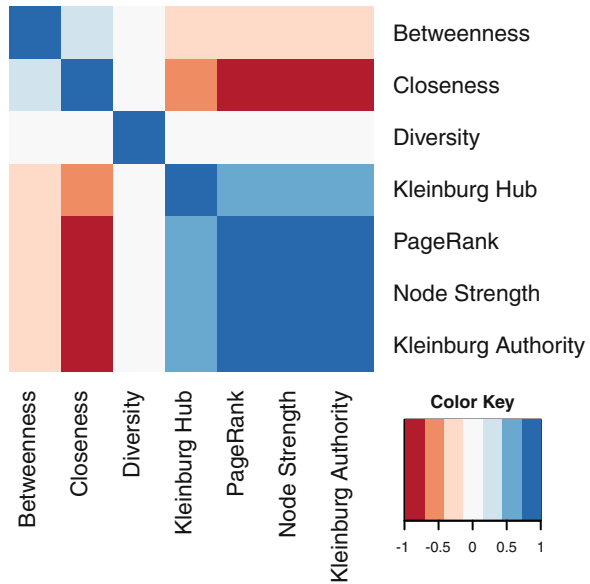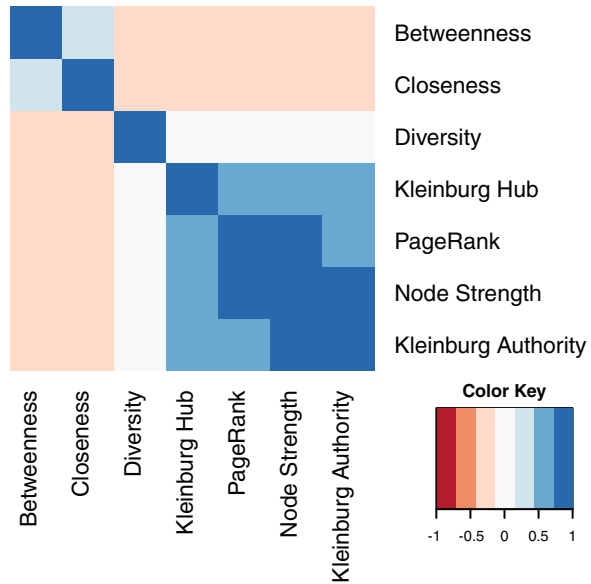
Fig. 2 Centrality measure correlations (τ values) for industries in the Waikato Region as a single network region

If we aggregate the ten TAs up to a single Waikato Region, we can compare centrality measure correlations over a much smaller set of industry-location pairs (106 nodes instead of 1060). This produces slightly weaker correlations, as shown in Fig. 2, but overall the pattern remains.

4.2 Comparing Multiplier and Centrality Measures

Comparing rankings for the different multipliers calculated, we find that the correlation between the Industry (Type I) and Household (Type II) disruption multipliers is $\tau = 0.21$. This matches the literature which shows that the inclusion of the household sector has a large impact on the results [14]. When we look at the Disruption multiplier, we find that this has a correlation of $\tau = 0.38$ with the Industry multiplier and $\tau = -0.01$ with the Household multiplier. This shows that the Disruption model, which simulates a disruption propagating through the IO network, is identifying different key industries to the existing IO multiplier analysis. This has implications for regional disruption planning.

Comparing the three multipliers with the network centrality measures (Fig. 3a), we find that the overall correlations are lower (-0.33 to 0.36) but that overall the eigenvector-based centralities and the overall industry-location strengths tend to match up with the traditional IO multipliers. This is not unexpected as they are both based on linear algebra matrix calculations that are mathematically similar, and the numerics agree with this. More interestingly the path-based measures (betweenness and closeness) are much more strongly correlated with the Disruption multipliers.
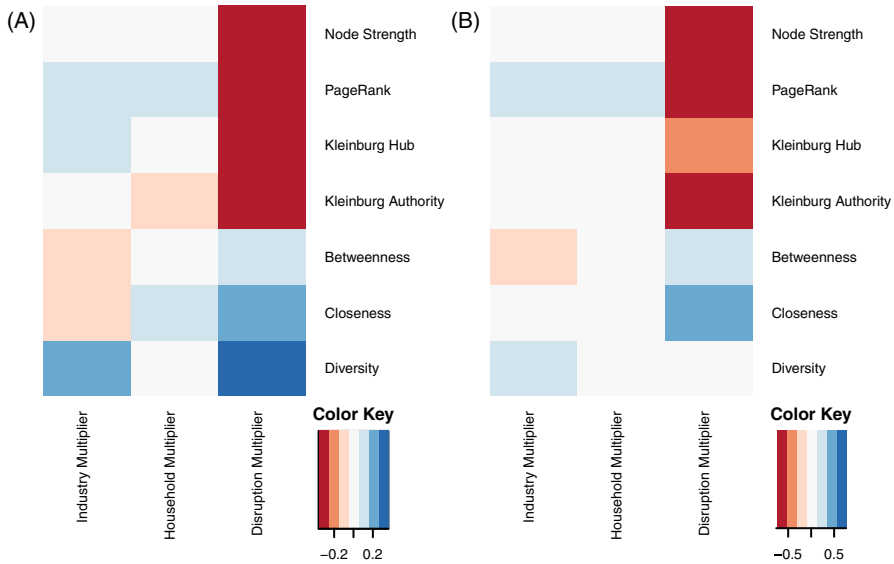
Fig. 3 Comparing the network centrality measures with: the two IO multipliers, the mean of the IO multipliers, and the disruption multiplier. This shows correlations between eigenvector-based centrality measures and IO multipliers, whereas path-based centrality measures correlate well with Disruption multipliers. There is a negative correlation between the two. For (**a**) industries in the ten TAs in the Waikato Region, and (**b**) industries in the Waikato Region as a whole

This makes intuitive sense as they are both concerned with flows and bottlenecks, that is with quantifying how disruptions to specific nodes flow on to impact the activity of dependent nodes. Another point to note is that we find the having a high diversity score is connected to having both a high Industry multiplier and a high Disruption multiplier. This highlights the potential importance of rarer industries within economic networks.

Aggregating to a single Waikato Region, we find much the same results (Fig. 3b), but with slightly weaker correlations (and anti-correlations). This is due to the disruption modelling becoming more homogenous in terms of industry-location distribution and activity when looking at the aggregated Region. A feature of the Disruption model is that it was designed to consider lower levels of aggregation, with the aim to be able to provide detailed results at single industry-location level resolution.

4.3   Impact of Spatial Aggregation

For all the analyses performed we considered the IO tables and economic networks with the Waikato Region broken down into 10 subregions (TAs) as well as with the whole Waikato Region considered together. This allowed us to look at the

impact of the spatial aggregation on the industries identified as as important from the disruption analysis.

We find that the eigenvector-based (strength-based) network measures identify the same key industries at both Region and TA levels of resolution. We find the same pattern for the Industry and Household multipliers. The value in analysing the system at the TA level disaggregation then becomes simply the ability to identify which TA the identified industry is most important to—it does not change which industries are identified. The exception is for industries that are disproportionately (or uniquely) represented in one or two TAs; for example, Coal Mining in the Waikato District, and Hospitals in Hamilton City. In these cases looking at TA level allows these to be ranked higher in importance than they would be if aggregating up to Regional level. Selected examples are given in Table 1.

However, for any path-based measures and for the Disruption multipliers, the level of spatial aggregation has a large impact on which industries are identified as important. Examples are given in Table 2. This can be explained as follows: aggregating the network up changes its structure—for example, at TA level there are fewer individual businesses within each industry categorisation, so the self-loops are smaller. Furthermore, the proportion of inter-industry flows that are within the TA itself is quite low (14–32%), with the majority of flows into (or out) of each industry coming from (or going to) other TAs within the Waikato Region and the rest of NZ. When considering the whole region, the proportion of inter-industry flows that stay within the region increases to around 60%. This is still far below the equivalent proportions that are typically observed in the literature when looking at IO networks at a whole country level [12]. It is therefore worth noting that metrics that are applicable for national level analysis may not behave as expected when working with disaggregated regional data, such as that considered here.

5   Discussion

In this work we have considered the question of how to identify industries that have a large impact on an economic system when they are disrupted. A goal of this paper was to show that network science measures can provide new useful tools for targeting interventions to reduce the impacts of disruptions on regional economies. In order to approach this, we have considered traditional IO multipliers, a new disruption model multiplier, and a range of network centrality measures. We have found that although traditional IO measures and eigenvector-based centrality measures are good at picking out the largest industries in terms of gross output or value-added, they do not match up with the industries that disruption modelling shows to have a large amplifying effect. We find that path-based measures, such as betweenness and closeness centrality, are far better at identifying industries that would have large flow-on impacts. These path-based methods explicitly consider the flow of money through the economy, and we find that the industries identified on these measures depend strongly on the level of spatial resolution.

Table 1 Selection of industries identified as important that are the same at Regional and TA level

Measure	Regional level	TA level
Total industry value added	Dairy cattle farming	Owner-occupied property operation (Hamilton City)
	Owner-occupied property operation	Hospitals (Hamilton City)
		Dairy cattle farming (Matamata-Piako)
Node strength	Dairy product manufacturing	Dairy product manufacturing (3 TAs)
	Dairy cattle farming	
	Electricity generation and on-selling	Electricity generation and on-selling (1 TA)
		Dairy cattle farming (1 TA)
		Coal mining (Waikato District)
PageRank	Dairy product manufacturing	Dairy product manufacturing (3 TAs)
	Dairy cattle farming	
	Electricity generation and on-selling	Meat and meat product manufacturing (1 TA)
		Dairy cattle farming (1 TA)
		Coal mining (Waikato District)
Industry Multiplier	Electricity generation and on-selling	Electricity generation and on-selling (6 TAs)
	Primary metal and metal product manufacturing	Primary metal and metal product manufacturing (7 TAs)
	Dairy product manufacturing	Dairy product manufacturing (7 TAs)
	Meat and meat product manufacturing	Meat and meat product manufacturing (7 TAs)
Household Multiplier	Preschool education	Preschool education (10 TAs)
	Postal and courier pick up and delivery services	Postal and courier pick up and delivery services (10 TAs)
	Specialised food retailing	Specialised food retailing (10 TAs)

In considering a natural hazard disruption, both the total size of the industry and the proportion its impact gets amplified by will play a role in determining the resulting impacts. By taking a network science approach, we are able to get a fuller picture of the potential targets for mitigation investment (e.g. stockpiling goods, having back-up generators in case of electricity outages).

In most disruption events, the impact will not be homogeneous through space. In most cases we would like to be able to consider the impact of a disruption on the well-being of communities, instead of just at national or even regional level. This is especially true for smaller localised events, which will not have a large impact at a national or regional level, but that could devastate a community. We have found that by considering smaller spatial units (in this case TA level) it is possible to get a better estimate of where the impacts will fall, as well as where susceptibilities are.

Table 2 Selection of industries identified as important on measures that differ Regional and TA level

Measure	Regional level	TA level
Betweenness	Petroleum and coal product manufacturing	Motor vehicle and motor vehicle parts wholesaling (Hauraki)
	Defence	Sewerage and drainage services (Thames-Coromandel)
	Sewerage and drainage services	Waste collection, treatment and disposal services (2 TAs)
	Air and space transport	Health and general insurance (Taupō)
		Petroleum and coal product manufacturing (Matamata-Piako)
		Warehousing and storage services (Waitomo)
Closeness	Air and space transport	Other transport (Otorohanga)
	Petroleum and coal product manufacturing	Motor vehicle and motor vehicle parts wholesaling (Hauraki)
	Beverage and tobacco product manufacturing	Warehousing and storage services (Waitomo)
	Warehousing and storage services	Health and general insurance (Taupō)
		Waste collection, treatment and disposal services
		Polymer product and rubber product manufacturing (2 TAs)
		Machinery manufacturing
		Electricity transmission and distribution
Disruption Multiplier	Defence	Other transport (Otorohanga)
	Life insurance	Motor vehicle and motor vehicle parts wholesaling (Hauraki)
	Petroleum and coal product manufacturing	Warehousing and storage services (Waitomo)
	Warehousing and storage services	Waste collection, treatment and disposal services (Waitomo)
		Polymer product and rubber product manufacturing (3 TAs)
		Health and general insurance (Taupō)

Even for the measures that do not change much between Region and sub-regional (TA) level (Table 1), looking at a higher granularity allows one to identify the unique (spatially specific) industries e.g. Hospitals and Coal Mining, that would be missed at a Regional level.

In future, increased data collection will make it possible to create networks at individual firm level. Making sure that we understand how different measures scale

from National to Regional all the way to individual firm level will be an important focus of future research.

References

1. Acemoglu, D., Carvalho, V.M., Ozdaglar, A., TahbazSalehi, A.: The network origins of aggregate fluctuations. Econometrica **80**, 1977–2016 (2012). https://doi.org/10.3982/ECTA9623
2. Blöchl, F., Theis, F.J., Vega-Redondo, F., Fisher, E.O.N.: Vertex centralities in input-output networks reveal the structure of modern economies. Phys. Rev. E **83**(4), 046127 (2011)
3. Cerina, F., Zhu, Z., Chessa, A., Riccaboni, M.: World input-output network. PLoS One **10**(7), e0134025 (2015)
4. Contreras, M.G.A., Fagiolo, G.: Propagation of economic shocks in input-output networks: a cross-country analysis. Phys. Rev. E **90**(6), 062812 (2014)
5. Csardi G., Nepusz T.: The igraph software package for complex network research. InterJ. Complex Syst. **1695** (2006). http://igraph.org
6. Dietzenbacher, E., Miller, R.E.: Reflections on the inoperability input-output model. Econ. Syst. Res. **27**(4), 478–486 (2015). https://doi.org/10.1080/09535314.2015.1052375
7. Harvey, E.P., Mead, S., Ayers, M.A., Smith, N.S., McDonald, G.M.: Quantifying the flow-on impacts from short-term natural hazard disruptions in local economies within New Zealand. (in preparation)
8. Hausmann, R., Hidalgo, C.: The network structure of economic output. J. Econ. Growth **16**(4), 309–342 (2011)
9. Hill, M.O.: Diversity and evenness: a unifying notation and its consequences. Ecology **54**, 427–432 (1973)
10. Kendall, M.: A new measure of rank correlation. Biometrika **30**(1–2), 8189 (1938). https://doi.org/10.1093/biomet/30.1-2.81
11. Liang, S., Qu, S., Xu, M.: Betweenness-based method to identify critical transmission sectors for supply chain environmental pressure mitigation. Environ. Sci. Technol. **50**, 1330–1337 (2016)
12. McNerney, J.: Network properties of economic input-output networks. IASA Interim Report. IIASA, Laxenburg (2009)
13. McNerney, J., Fath, B.D., Silverberg, G.: Network structure of inter-industry flows. Phys. A Stat. Mech. Appl. **392**(24), 6427–6441 (2013)
14. Miller, R.E., Blair, P.D.: Input-Output Analysis Foundations and Extensions, 2nd edn. Cambridge University Press, Cambridge (2009)
15. Oosterhaven, J.: On the limited usability of the inoperability IO model. Econ. Syst. Res. **29**(3), 452–461 (2017). https://doi.org/10.1080/09535314.2017.1301395
16. Xu, M., Allenby, B.R., Crittenden, J.C.: Interconnectedness and resilience of the US economy. Adv. Complex Syst. 14(5), 649–672 (2011)