Economic Systems Research

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/cesr20

Filter methods for MRIO tables: an evaluation

Markus Simbürger

ECONOMIC SYSTEMS RESEARCH
https://doi.org/10.1080/09535314.2021.1971629

Filter methods for MRIO tables: an evaluation

Markus Simbürger[a,b]

[a]Joanneum Research – LIFE, Graz, Austria; [b]Institute of Economics, University of Graz, Graz, Austria

ABSTRACT

Researchers who deal with network analysis based on multi-regional input–output (MRIO) tables cannot avoid the intensively discussed issue of filtering, which means identification of the most important and significant trade connections. The question of what is an appropriate filter method remains. This paper expands the existing discussion and brings new insight based on the evaluation of existing filter methods for MRIO tables. Six filter methods from the prevailing literature are identified as relevant and tested on the published MRIO tables: EORA26 and EXIOBASE. The results are verified by a case study. The evaluation shows that the Tolerable Limit approach and filter based on the Weaver–Thomas Index are the most restrictive. The Leontief filter and the filter based on holistic accuracy can be partially recommended. The filter on absolute trade values and average transactions can be recommended as 'good' methods.

ARTICLE HISTORY
Received 16 December 2020
In final form 19 August 2021

KEYWORDS
Filter methods; important trade connections; MRIO tables; input–output analysis

1. Introduction

The analysis as well as the construction of input–output (IO) tables has improved over time due to better data availability and increased computation power. Today it is possible to construct IO not only for single countries but also for multiple regions–multi-regional input output (MRIO) tables. When MRIO tables are rigorously investigated, one can find the presence of values smaller than a unit of currency. Clearly, they embody little economic meaning. They exist, however, because they are compiled via various data sources. Moreover, economies have different structures and statistical agencies use different base sectoral definitions, harmonize their own codes in distinct manners, and aggregate further when pulling data together in the process of producing the tables (Lenzen et al., 2013; Stadler et al., 2018). In addition, trade is reported asymmetrically between countries. That is, the global sum of exports does not even equal the global sum of imports (Gaulier & Zignago, 2010); so, reconciliation and balancing approaches must be applied to balance them. This induces those small trade values. While the small values are economically unimportant, they play a large role in network analysis because each trade flow, independent of how small this trade flow might be, identifies that the two sectors are linked. With this in mind, the operation of filtering can play a role. Filtering defines values under a certain threshold as unimportant or insignificant and thereby assigns a value of zero to them. That

CONTACT Markus Simbürger markus.simbuerger@joanneum.at; markus.simbuerger@edu.uni-graz.at

is, filtering identifies the most important connections. A structural analysis of the resulting reduced network readily enables analysts to find groups of connected industries or sectors or allows us to distill how industries or sectors interact in the economy (e.g. Aroche-Reyes, 2001; Chopra & Khanna, 2015; Okamoto, 2005; Xiao et al., 2019; Zhang et al., 2016). Furthermore, those networks can be further analyzed to identify, for example, systemic risks (Acemoglu et al., 2012, 2016; Bierkandt et al., 2014; Wenz et al., 2014). So far, scientists have applied filter methods without empirical verification of the quality of the method used; furthermore, filtering is only one of many process steps and, therefore, not typically discussed explicitly in much detail. In writing this paper, I try to change this. The aim of this paper is twofold; first, it summarizes discussions on filtering, broadly reviews existing filter methods, and explains their differences. Second (and the main contribution), I evaluate selected filter methods by applying them to two published MRIO tables and, hopefully, present some guidance to future applications of filters. Although the applications focus on MRIO tables, findings should also be relevant to all IO tables. Thanks to readily available computational power, I can investigate filter methods in this way. I show that filtering is relevant to ongoing network analyses based on IO and MRIO tables with the hope of encouraging others to think carefully about the filter methods they apply.

2. Literature review

As mentioned in the introduction, filtering concerns the process of identifying the relevant or most important trade connections. With the elaboration of MRIO tables, the purpose of filtering changed. The purpose is not only the identification of the relevant trade connections but also comprises the identification of realistic trade connections. This means that the procedure of building IO tables leads to statistical inaccuracy of the estimated values; therefore, in the context of filtering, terms like 'sensitivity', 'accuracy' or 'tolerance' also pop up in the literature. Historically, the first attempts to indicate the most relevant connections in an economy were applied by Leontief in 1965 with the purpose of identifying principal suppliers to certain industries of the U.S. Economy (Leontief, 1965).

Paelinck et al. (1965) originally conceived what has become known as hypothetical extraction method (HEM) to identify the most important sectors in an economy. This method was first employed in English by Schultz (1976) who cited Strassert (1968).[1] HEM extracts one sector at a time and, thereby, can provide an index of the influence of each sector on an economy. Separately, Miller (1966, 1969) applied the same approach to examine the effects on a national economy of extracting a region, rather than a sector. Among others, Dietzenbacher et al. (1993) applied the method to an international scale in order to identify important linkages of national economies within a group of nations.

Jensen and West (1980) noted that few researchers had investigated the analytical sensitivity of IO components.[2] This topic focuses on how error or reliability with respect to coefficients (accuracy) affects economic outcomes. Jensen and West (1980) identified that small coefficients have less effect on output and related multipliers, and that the size of that effect rises with coefficient size. With this in mind, the authors defined four types of accuracy: (i) A-Type accuracy refers to the degree to which an IO table represents the 'true

[1] An overview of the method and further literature supplements are provided by Miller and Lahr (2001)
[2] For a more detailed discussion on important coefficients and sensitivity see Thakur (2011).

table' for the economy. It is only possible to try to come close to the 'true table', however, because two sources of error can be identified: data measurement error and error inherent to table compilation. (ii) B-Type accuracy refers to the exactness of an input–output model and the reflection of the operation of an economy. (iii) Partitive accuracy focuses on the cells of the table and focusses on cell-by-cell accuracy. If each cell is accurate and records the 'true' transactions, then the table as a whole should better reflect the 'true table'. (iv) Holistic accuracy, instead, emphasizes the table's ability to be accurate at a broad scale. This suggests that cellwise importance is less critical and that the focus should be on proper representation of the main features of the economy (Jensen, 1980).

In a similar way to Jensen and West (1980) and Songling and Gould (1991) considered the importance of coefficients in the context of error transmission and coefficient change. The authors investigated in more detail how sensitively the final demand, output multipliers, or gross output react to a change in the coefficients. Important coefficients are those that play a major role in error transmission (Songling & Gould, 1991).[3]

Cassetti (1995) proposed a new algorithm to identify important coefficients (ICs) by avoiding the choice of arbitrary thresholds. The algorithm identifies the most significant intermediate transactions by means of a 'representative index', where groups of sectors are defined by high interdependence and by triangularization of the structure. For this index, the calculation of several Leontief inverses is necessary. Depending on the size of the matrix, as mentioned by the authors, a middle-sized IO table would already require the calculation of millions of inverses. For MRIO tables, calculation of the number of inverses would increase dramatically and require high computation power and time, which currently renders this algorithm impractical in some cases without access to a supercomputer[4].

Aroche-Reyes (1996) also tried to avoid arbitrary thresholds and use only endogenous information. The author applied a reformulated algorithm by Jilek (1971) and Schintke and Stäglin (1988), which yields tolerable limits for each technical coefficient. The algorithm does not measure how much the output changes if the coefficient varies by a certain amount. By contrast, it measures how much the coefficient has to change in order to alter the output by 1%. This sets a tolerable limit (TL) for the coefficients. Tarancón et al. (2008) compared the tolerable limit approach with alternative approaches, based on an 'elasticity' concept or linear programming to solve a system of equations. The authors admit that, from a practical perspective, the tolerable limit approach leads to similar results compared to alternatives; however, if the aim is to identify ICs according to social welfare, which means in the context of impact on the whole economy, then the ICs on the elasticity would be preferable. Linear programming is to be preferred when more complex scenarios are employed.

Without a detailed discussion on the filter method itself, Okamoto (2005) applied a hybrid approach to investigate the agglomeration of intraregional and interregional linkages in China. The applied approach is based on the average transaction amount.

A new approach was implemented by Luu et al. (2017). They used a hypergeometric method, built on probabilities of trade connections, to filter data. After applying this

[3] A good overview of important coefficients is given by Miller and Blair (2009).
[4] For example see Lenzen (2019).

method, only those links remain which are sensibly high with respect to randomly chosen connections. This method provides a significance level, which can be tested against the null-hypothesis that the connections are a random co-occurrence. The method goes back to Tumminello et al. (2011); however, it has the drawback that the underlying system has to be bipartite and is, therefore, only applied to evaluate trade connections between countries. The trade connections within a country between the different sectors cannot be evaluated. Recently, studies by Xu and Liu (2013), Zhang et al. (2016) or Xiao et al. (2019) applied the Weaver–Thomas Index, which is based on the distribution of the trade flows, for filtering.

Although the focus here is on quantitative studies, I should also briefly mention qualitative studies. For example, Holub et al. (1985) focus on qualitative aspects of IO analysis. Such qualitative analysis derives an adjacency matrix that represents sector's direct delivery paths. A direct delivery path is indicated when the original value of the transaction exceeds a certain 'bagatelle' amount (BAG-value). With a single BAG-value, however, only a snapshot of the qualitative information can be shown. Thus, many analysts progressively raise the BAG-value to show transformations in the resulting adjacency matrix (Holub et al., 1985). Schnabl (1994) developed the minimum flow analysis (MFA) based on such a multi-layer procedure. de Mesnard (2001) criticized it arguing that the different layers of the Boolean matrix required in MFA are always the same: therefore, the added layers provide no additional information.

The above gives a good anchor point for interested readers. The next section describes the selected filter methods in more detail.

3. Filter methods

As mentioned in the introduction, filtering defines values under a certain threshold as unimportant or insignificant. The decision criteria for thresholds can take various forms, as can the threshold values. Some filter methods are not applicable for the current research question or are too computationally demanding when applied to some larger MRIO tables. Given these caveats, I deemed the following to be the most relevant filter methods: Leontief (1965)'s (LEO), a tolerable limit approach (TLA) as used by Aroche-Reyes (1996), one based on Jensen and West (1980)'s concept of holistic accuracy (HA), another on the average transaction amount (AT) as applied by Okamoto (2005), and one based on the Weaver–Thomas Index (WTI). Not mentioned explicitly in the literature review but also relevant is a filter based on absolute trade values (TV). For simplicity, the filter methods without indices for regions. So, the indices i and j refer only to sectors where sector i is delivering a certain amount z to sector j.

3.1. *Filtering on absolute trade values*

The simplest and most obvious filter method is built on absolute intermediate interindustry trade values denoted by z_{ij}. The function of the threshold value represents then an absolute value f^{abs} such that:

$$z_{ij}^{f} = \begin{cases} 0, & \text{if } z_{ij} < f^{abs} \\ z_{ij}, & \text{otherwise} \end{cases} \tag{1}$$

The literature yields no recommendations concerning a proper absolute threshold. But examples of applied thresholds exist.

The *Australian Bureau of Statistics* states that, due to modeling techniques, the statistical accuracy of relatively small values cannot be sufficiently verified. They mention thresholds from AU \$ 500,000 to AU \$ 1 million, below that values carry little, if any, economic meaning (ABS, 2017, 2019). Also, Bierkandt et al. (2014) and Wenz et al. (2014) note that small values can spread damage in a seemingly quirky, diluted manner; so they apply a threshold of 1 million US \$ and neglect all smaller flows (Bierkandt et al., 2014; Wenz et al., 2014). COMTRADE uses a minimum 1000 US \$ as a threshold (Gaulier & Zignago, 2010). Since COMTRADE data or bilateral balanced trade data derived from them (BACI database) are used to build MRIO tables, a threshold of 1000 US \$ appears most relevant to the research reported in this paper.

3.2. Leontief

While the father of IO analysis, Leontief also devised the first filtering method. It is not based on absolute values of shipments between sector j and i but rather on the values of technology coefficients a_{ij}. These coefficients are obtained by dividing the value of the IO table z_{ij} by the total input of the sector x_j; hence, they denote the fixed relationship between outputs and required vectors of inputs of sectors. The filter criteria applied by Leontief are:

$$a_{ij}^{Leo} = \begin{cases} 0, & a_{ij} < f^{leo} \text{ where } f^{leo} = \dfrac{1}{n} \\ a_{ij}, & \text{otherwise} \end{cases} \tag{2}$$

where n denotes the number of sectors in the relevant IO table.

This filter method has been applied, for example, by Aroche-Reyes (2001) to identify strongly connected components (SCCs). The purpose of filtering is to gain more insight into the structure of a very dense matrix by making it sparser. Aroche-Reyes (2001) further adjusted Leontief's filter by altering the filter threshold, such that $f^{leo} = 1/2n$ or $f^{leo} = 1/3n$. Such thresholds produce equivalent but not necessarily identical results to the SCCs (Aroche-Reyes, 2001).

3.3. Tolerable limits approach

The tolerable limits approach (TLA) measures the importance of technology coefficients a_{ij}. Those with a large impact on the output of many sectors are deemed important coefficients (ICs). TLA identifies the degree to which a technology coefficient a_{ij} must change in order to alter output by, at most, by 1%, given fixed final demand. The tolerable limits r_{ij} are calculated:

$$r_{ij} = \frac{1}{a_{ij}[\alpha_{ji} + (\alpha_{ii}/x_i)x_j]} \tag{3}$$

where α_{ij} denotes the entry in the Leontief inverse[5]. The variables τ_i and τ_j denote the sectoral gross output values. Very sensitive coefficients, with a value of r_{ij} smaller than 20%

[5] For the Leontief inverse matrix $(\mathbf{I} - \mathbf{A})^{-1}$, the entries a_{ij} of the matrix $\mathbf{A}$ are calculated as $a_{ij} = z_{ij}/x_j$, where x_j is the total output of sector j.

($r_{ij} < 0.2$), are indicated as ICs (Aroche-Reyes, 1996; Jilek, 1971; Schintke & Stäglin, 1988).

$$a_{ij}^{TLA} = \begin{cases} 0, & r_{ij} > f^{TLA} \text{ where } f^{TLA} = 0.2 \\ a_{ij}, & \text{otherwise} \end{cases} \tag{4}$$

ICs are technology coefficients for which longer sequences of indirect connections or larger sets of sequences are prevalent. In other words, ICs are identified when two directly connected sectors are also indirectly connected to many other sectors. Furthermore, this means that if an IC of a_{ij} is important to sector i, but sector i itself is less important to the broader economy than the IC a_{ij} must also be less important.

3.4. Holistic accuracy

Holistic accuracy focuses on the general representativeness of an estimated IO table and its ability to capture the synergistic characteristics of the economy it depicts. Here, the accuracy of any given single element is secondary at best, as long as model results show a realistic picture. Holistic accuracy essentially measures the effect of relative coefficient size on IO multipliers. It basically assumes there exists some threshold for technology coefficients that accounts for multiplier size rather than for the size of particular interindustry interactions as in TLA.

For the investigation of an appropriate threshold, the cumulative effect of the relative coefficient size is of interest; therefore, the coefficients a_{ij} and output multipliers $m(0)$ are derived from the original IO table $\mathbf{Z}$. The 'true' multipliers $m(0)$ are the reference. The output multiplier is calculated in the conventional way, whereby entries of Leontief inverse $(\mathbf{I} - \mathbf{A})^{-1}$ are summed over the column of the sector i (Richardson, 1972).

For the calculation of the cumulative effect, the reduced matrix $\mathbf{Z}(\mathbf{s})$ with the elements $z_{ij}(s)$ with a set of thresholds $a(s), s \in \{1, \ldots, n\}$ is derived, such that:

$$z_{ij}(s) = \begin{cases} 0, & \text{if } a_{ij} < a(s) \\ z_{ij}, & \text{otherwise} \end{cases} \tag{5}$$

As thresholds $a(s)$ it is recommended to use $n = 20$ and the $\frac{s}{n}$-quantile of non-zero coefficients ($a_{ij} > 0$) (Jensen & West, 1980). After each removing step, the multipliers $m(s)$ are again calculated from the reduced matrix $\mathbf{Z}(\mathbf{s})$. Based on this analysis, the filter threshold $a(s^*)$ could be set on the 1% reduction of the input–output multiplier, such that the following condition is satisfied: $s^* = \max\{s : \frac{|m(0)-m(s)|}{m(0)} < 1\%\}$. The pioneering work by Jensen and West (1980) shows that a large number of small elements in an IO table can be removed before a significant effect occurs.

3.5. Filtering on the average transaction amount

Okamoto (2005) applies a hybrid approach based on the absolute trade values and the information of the technological coefficients of the Leontief inverse and derives the average transaction amount for his investigation of the agglomeration of intraregional and interregional linkages in China. The filter is:

$$z_{ij}^{average} = \begin{cases} 0, & \text{if } z_{ij} < \alpha \\ z_{ij}, & \text{otherwise} \end{cases} \tag{6}$$

where the average transaction amount $\alpha = i'(\mathbf{L} - \mathbf{I})i/n^2$, i is a unit vector of the length n (dimension of the MRIO table) and $\mathbf{L}$ is the Leontief Inverse matrix $(\mathbf{I} - \mathbf{A})^{-1}$. As Okamoto (2005) highlights, the filter value should be carefully selected; therefore, the number of intermediate transactions that are greater than α is compared to the number of intermediate transactions that are 10, 100, or 200 times larger than the average transaction. The final filter value chosen by the authors is ten times the size of an average transaction.

3.6. Weaver–Thomas index

Xu and Liu (2013) and Xiao et al. (2019) use the Weaver–Thomas Index in IO-related analysis and network analysis. Basically, this index compares an observed distribution with an assumed one by calculating and comparing the quadratic sum. The threshold value is identified by the closest approximation of the distributions and shows how many values have to be selected in order to be as close as possible to the assumed distribution. The formula is:

$$\omega(l) = \sum_{k=1}^{n^2} \left[s(k,l) - 100 \times \frac{a_k^\star}{\sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij}} \right]^2, \tag{7}$$

where the term $s(k,l)$ represents the assumed distribution such that

$$s(k,l) = \begin{cases} \dfrac{100}{l} & (k \leq l) \\ 0 & (k > l) \end{cases} \tag{8}$$

and the observed one is represented by the second term in Equation 7. In the case of IO tables, a matrix $\mathbf{A}$ with n sectors is used. All elements of $\mathbf{A}$ are vectorwise stacked in descending order to create vector $A^\star$, where l denotes the lth element of $A^\star$ and $a_k^\star$ is the kth element of vector $A^\star$. The threshold value $a^\star$ is the $l^\star$th element of $A^\star$, where $\omega(l^\star) = min\{\omega(1), \omega(2), \ldots, \omega(n \times n)\}$ (Xiao et al., 2019; Xu & Liu, 2013; Zhang et al., 2016). The filtered matrix is derived as:

$$a_{ij}^{WTI} = \begin{cases} 0, & \text{if } a_{ij} \leq a^\star \\ a_{ij}, & \text{otherwise} \end{cases} \tag{9}$$

To apply the Weaver–Thomas index to MRIO tables, n does not only represent the sectors but all sector-country combinations.

4. Numerical comparison of results

As mentioned earlier, filtering not only identifies the most relevant or important trade connections but also those interindustry connections that have been measured most reliably.

Investigating the effect of the filter methods on MRIO tables (e.g. EXIOBASE, EORA26) enables comparisons of the filtered and unfiltered matrices to determine how well the former represent the latter. The investigation is carried out for two different types of MRIO tables. An important aspect of my analysis of the MRIO tables is that both the number of regions and the number of sectors differ across them. This allows the value of shipments to differ between them even if the aggregated economies are identical. For example, the WIOD database (Timmer et al., 2015) represents agriculture, hunting, forestry, and fishing products with a single sector, whereas EXIOBASE represents them with 19 different sectors. That is, the EXIOBASE database subdivides the same trade values into several sectors so that values of shipments are necessarily everywhere smaller for those industries compared those in WIOD. When comparing the tables, such different country and sector structures should be taken into account. For this reason, I opted to investigate filter methods using two different sets of MRIO tables.

The two different MRIO tables are the EORA26 and EXIOBASE.[6] EORA26 has a more or less 'standard' sectoral resolution but extreme detail in the number of countries, a few of which have never produced their own IO tables. EXIOBASE contains many more sectors, particularly including more related to agriculture, which is not usual in 'standard' IO resolution but far fewer countries. These very different underlying structures enable an analysis of results of filter methods that might emanate from such structural differences.

EORA26 (Lenzen et al., 2012, 2013) is available for years from 1990 to 2015 and contains 26 sectors and 190 countries with one rest of the world (ROW) sector, which includes just one sector and is used to account for the statistical discrepancies. The tables are available at both basic and purchaser prices. I use the tables in basic prices. EXIOBASE Stadler et al. (2018) provides data for years from 1995 to 2016 in current basic prices. I use the product-by-product monetary EXIOBASE3 database. It includes 44 countries, 5 Rest of the World regions, and 200 products among other data items. For the investigation, I use data for 2011.[7]

4.1. Measurement criteria for filter methods

I apply three measures to identify the quality of the six filter methods. They are (i) the relative number of trade connections, (ii) standardized total percentage error (STPE), and (iii) a modified value of the weighted absolute distance (WAD). The number of trade connections or links (network terminology) is a fundamental property. To understand the comparability of results, the number of trade connections (links) are stated in relative terms to the total number of trade connections without filtering.

The second criterion, the standardized total percentage error (STPE) by Miller and Blair (1983), also applied under another name by Sawyer and Miller (1983) and Szyrmer (1984), represents the mean absolute error divided by the mean value of the unfiltered matrix. This is shown in Equation 10, where z_{ij} represents the values of the unfiltered matrix and $\hat{z}_{ij}$ of the filtered matrix. To put it simply, the STPE ranges from zero to one and measures, in our case, the relative trade volume represented by the filtered matrix. For interpretation purposes, STPE* is calculated such that a value of one means that a hundred

[6] EXIOBASE data are available for free, although registration is mandatory. EORA26 data also are freely available for academic use at degree-granting academic institutions; other users must license the data.

[7] When I started this research, only data through 2011 were available.

percent of the trade values are represented by the filtered matrix, and zero that no trade values are represented by the filtered matrix.

$$\text{STPE}^* = 1 - \frac{1/n^2 \sum_j \sum_i \mid z_{ij} - \hat{z}_{ij} \mid}{1/n^2 \sum_j \sum_i z_{ij}} = 1 - \frac{\sum_j \sum_i \mid z_{ij} - \hat{z}_{ij} \mid}{\sum_j \sum_i z_{ij}} \tag{10}$$

The STPE* is an important criterion because the aim of MRIO tables is to cover total trade flows as well as possible and, furthermore, global GDP. The only drawback of the STPE is that the measurement is not sensitive to high-value cells (Lahr, 2001); therefore, the third criterion is introduced.

The third criterion is based on a proposed matrix difference measure by Lahr (2001) and is called weighted absolute distance (WAD); this is represented in Equation 11.

$$\text{WAD} = \frac{\sum_j \sum_i (z_{ij} + \hat{z}_{ij}) \mid z_{ij} - \hat{z}_{ij} \mid}{\sum_j \sum_i (z_{ij} - \hat{z}_{ij})} \tag{11}$$

The term $(z_{ij} + \hat{z}_{ij})$, in Equation 11, weights the absolute difference $\mid z_{ij} - \hat{z}_{ij} \mid$ such that the errors of large cells are emphasized (Lahr, 2001). A drawback of WAD is that its range depends on the size of the referent matrix. Thus, WADs of EORA26 and EXIOBASE would not be comparable. So, to enable comparisons, I normalized WAD. I did this by rescaling WAD values relative to its maximum value and multiplying by 100. Such a normalized WAD, named WAD^{rel}, ranges from zero to 100. The maximum of the WAD is, if all elements $\hat{z}_{ij}$ of the compared matrix are equal to zero:

$$\max(\text{WAD}) = \frac{\sum_j \sum_i (z_{ij}) \mid z_{ij} \mid}{\sum_j \sum_i (z_{ij})} \quad \text{where } \hat{z}_{ij} = 0 \tag{12}$$

From this, the relative weighted absolute distance can be defined:

$$\text{WAD}^{rel} = \frac{\text{WAD}}{\max(\text{WAD})} * 100 \tag{13}$$

such that $0 \leq \text{WAD}^{rel} \leq 100$. The newly introduced measure identifies how far away the actual weighted absolute distance is from the maximum weighted absolute distance. This enables comparisons of matrices of different dimensions, the case for EORA and EXIOBASE.

4.2. Results and discussion

The results of the analysis are summarized in Table 1, and are discussed in Section 4.2.1 to 4.2.6. The first column indicates the filter method that was applied. The second and sixth column show the filter value by the corresponding database, respectively EORA26 and EXIOBASE. Note that the filter values are quite different by a method of the holistic accuracy (HA), the Leontief filter (LEO), the filter based on the Weaver–Thomas index (WTI) and for the filter on average transaction amount (AT).

The other columns show the relative number of trade connections (Rel.Nr.Lins), the standardized percentage error (STPE*), and the relative weighted absolute distance (WAD^{rel}) for the EORA26 and EXIOBASE separately.

Table 1. Results for the filter method of EORA and EXIOBASE.

	EORA26				*EXIOBASE*			
Filter	Filter value	Rel. Nr.Links	STPE*	WAD^{rel}	Filter value	Rel. Nr.Links	STPE*	WAD^{rel}
TV	0.10	0.9646	1.000	0.000	0.10	0.4826	1.000	0.000
TV	1.00	0.3986	1.000	0.000	1.00	0.3174	1.000	0.000
TV	500.00	0.0241	0.998	0.000	500.00	0.0367	0.992	0.000
TV	1000.00	0.0183	0.996	0.000	1000.00	0.0256	0.987	0.000
LEO	1/3n	0.0015	0.755	0.489	1/3n	0.0091	0.869	0.062
LEO	1/2n	0.0011	0.692	1.064	1/2n	0.0069	0.840	0.124
LEO	1/n	0.0006	0.578	2.231	1/n	0.0042	0.774	0.402
WTI	0.0452	0.0005	0.552	2.582	0.00949	0.0027	0.703	0.880
TLA	1.000 %	0.0070	0.937	0.005	1.000 %	0.0126	0.916	0.014
TLA	100 %	0.0018	0.814	0.156	100 %	0.0027	0.732	0.503
TLA	20 %	0.0005	0.529	4.146	20 %	0.0006	0.455	7.449
TLA	10 %	0.0002	0.371	11.516	10 %	0.0003	0.325	16.015
HA	9.86e−06	0.2167	0.997	0.000	3.33e−05	0.1000	0.987	0.000
AT	Θ	0.0961	1.000	0.000	Θ	0.0520	0.995	0.000
AT	$10 \times \Theta$	0.0387	0.999	0.000	$10 \times \Theta$	0.0156	0.978	0.000
AT	$100 \times \Theta$	0.0160	0.996	0.000	$100 \times \Theta$	0.0038	0.925	0.002
AT	$200 \times \Theta$	0.0118	0.993	0.000	$200 \times \Theta$	0.0024	0.896	0.005

Notes: TV = filter on absolute trade value, LEO = Leontief filter, WTI = Weaver–Thomas Index, TLA = Tolerable Limit approach, HA = holistic accuracy, AT = filter on average transaction amount. The filter value for the TV is thousand units. Total number of trade connections (links) for the EORA26 and EXIOBSE are 24,137,750 and 29,775,785, respectively. The total trade value for the EORA26 and EXIOBSE are 8.08e+10 (thousand $) and 5.01e+07 (million EUR), respectively. The number of sectors *n* for the EORA26 and EXIOBSE are 26 and 184, respectively. The filter value Θ for the AV for the EORA26 and EXIOBASE are 13.82 and 0.2439 units, respectively.

4.2.1. Filter on absolute trade values (TV)

The filter on trade values (indicated by TV in Table 1) shows the expected result for both matrices that the number of links decreases dramatically with higher filter values; however, the trade volume, shown in relative terms to the total trade volume, does not decrease much. This suggests that MRIO tables are inflated by many small elements. According to Wenz et al. (2015), flows that are smaller than 1 million US $ from the EORA26 table contribute, in total, less than 0.6% of the total value of shipments depicted by the database.

More insight into the filter value can be gained by a comparison of the results of the relative number of links for EORA26 and EXIOBASE. When the filter value is set to 0.10 (100 units), links remaining in EXIOBASE are fewer than those remaining in EORA26. This finding that the EXIOBASE is more sensitive to lower threshold values is in line with expectations because EXIOBASE is more disaggregated and, hence, must necessarily have a far greater share of small entries. It is remarkable that with a filter value of 1000 (1,000,000 units) more links remain in EXIOBASE than in EORA26. The reason for the converse result is the number of countries covered by the two MRIO tables. Recall, EORA26 includes 190 countries, whereas the EXIOBASE only covers 44 plus 5 more 'Rest of the World' regions. This means that international trade flows between the sectors are far more disaggregated in the EORA26; so, within these international flows, EORA26 has fewer values more than 1 million units than does EXIOBASE.

The measure of the relative weighted absolute distance (WAD^{rel}) suggests there are few large differences between the unfiltered MRIO tables and their filtered alternates. This suggests only that high-valued elements are preserved in the filtered version, confirming expectations.

The filter for absolute trade values must be, perhaps, more arbitrarily determined. This is because the literature informs us little about its 'best' threshold values, as mentioned in Section 3.1. An advantage of this approach is that it is easy to intuit what 1000 or 1 million units means. But sectors size plays a stronger role in this sort of filter. Thus, interindustry flows of the chemical sector can be more important than those of the agricultural sector simply because the chemical industry generates more output. In this vein, sectors are unequally penalized by the absolute filter due to their unequal economic might.

4.2.2. Leontief filter (LEO)

The results for the proposed Leontief filter are indicated by 'LEO' in Table 1. The first notable result is that varying the threshold from $1/n$ to $1/2n$ to $1/3n$ does not much change the overall results. That is, only a few trade connections remain and also the count of non-zero trade values is significantly reduced. With the Leontief filter, higher interindustry trade values are defined as unimportant if they do not contribute much to a sector relative to its output. That is, it is only the relative size of a sector's inputs to its output that matters. This yields an advantage of industries that supply smaller sectors and a disadvantage to those that ship inputs to larger sectors.

It should be mentioned that the filter values differ for the EORA26 and EXIOBASE matrix due to the number of sectors in each. This forces a filter value of $1/n$ for EXIOBASE to be lower than that for EORA26. Note that results show for the same filter value that the number of remaining links and covered trade values is larger for the EXIOBASE than for the EORA26. Furthermore, the WAD^{rel} is much higher for the EORA26 than for the EXIOBASE. This suggests that more large interindustry connections are identified as 'unimportant'. In this regard, when a single sector is disaggregated, the structure of all other sectors remains fixed. Still, n increases, and the filter threshold decreases. So, the Leontief filter adjusts to a more refined threshold value. But, in the end, it becomes less restrictive on sectors that were not disaggregated. Thus, it follows that when sectors are disaggregated, the filter changes to favor sectors that were not disaggregated. Of course, the converse is true; if sectors are aggregated, the filter changes to favor the aggregated sectors. Campbell (1975) highlighted the advantage of the Leontief's filter property of being endogenously defined. But, like others, it too is not devoid of quasi-arbitrary determinations; one must decide whether n, $2n$, xn or some other value is an appropriate denominator.

4.2.3. Weaver–Thomas index (WTI)

Like Leontief's filter, the WTI is also based on coefficients. But, the WTI compares the assumed and actual distribution of the coefficients, not only the number of sectors. The results show that the identified WTI for the EORA26 and EXIOBASE is lower than that for Leontief's filter. For this reason, the remaining links and covered trade values are also lower and the STPE is higher, compared those for Leontief's filter. The comparison of EORA and EXIOBASE shows that the WTI of EXIOBASE is much smaller than those for EORA26. This highlights again the measure's sensitivity to the number of sectors since its threshold lowers.

4.2.4. Tolerable limit approach (TLA)

The results in Table 1 show that the recommended value of 20% is much more restrictive than the filter for absolute trade values. Even increasing the threshold to a value of 1.000%

does not deliver results close to the TV filter. The reason is that in MRIO tables, the size of the effect on output of a change in a single technology coefficient tends to be very small compared to the magnitude of the output. The results of the TLA are closer to those of Leontief's filter.

The results are much clearer if the result for the number of trade connections (Rel.Nr.Links) is compared between EORA26 and EXIOBASE in Table 1. For each filter value, the number of trade connections is higher for EXIOBASE compared to that for EORA. On the one hand, the greater sectoral detail should lower the impact on total output. On the other, recall that some regions are subsumed within EXIOBASE's five Rest of the World regions, which raises the impact on the output. Because the impact from technology coefficients is more imbalanced in EXIOBASE, more trade connections are identified as relevant by TLA. Another aspect is that the relative number of trade values, indicated by the STPE*, is higher for EORA26 than for the EXIOBASE, although the number of remaining links is lower in EORA26. This indicates, furthermore, the imbalanced aspect of EXIOBASE, where trade connections with smaller values are indicated as economically relevant whereas trade connections with higher trade values are indicated as unimportant. Among all filter methods, the TLA has the highest values of the WAD^{rel}. But it should be noted that the WAD^{rel} ranges from 0 to 100, where a value of 10 should be viewed as comparatively small.

4.2.5. Holistic accuracy (HA)

A special issue for the filter method based on holistic accuracy is that for each IO table an individual filter value is identified instead of one overall filter value applied to both databases. In contrast to the separation of the trade values into 20 intervals proposed by Jensen and West (1980), I separated the matrices into smaller intervals,[8] up to 60, in order to find a threshold value such that the output multiplier is reduced by around 1% with a tolerance of ±0.2%.

This method's filter is also based on technology coefficients; so, it is remarkable that in compared to LEO and WTI, its filter value for EORA26 is lower than the equivalent for EXIOBASE.

For the EORA26 78% and for the EXIOBASE, 90% of the trade connections can be removed before the output multiplier reduces to around 1%. This contrasts starkly with results in Jensen and West (1980), who show that 30–35% of the smaller coefficients are removed before the output multiplier reduces by 1%. Using their filter value, resulted in more than 98% of the total trade values still available.

A comparison with the filter on absolute trade value in Table 1 shows that the filter method for holistic accuracy identifies more trade connections are relevant although fewer trade values remain. For example, the filter method TV with a filter value of 500 indicates 2.4% trade connections, which represents 99.8% of the trade values, whereas the filer method HA indicates more trade connections (21.7%) which represents fewer trad values (99.7%).

I should mention that the calculations required for this filtering method are quite time-intensive. For each interval, I had to calculate the Leontief inverse to test how much the

[8] The chosen procedure for the intervals is based on quantiles, whereas the intervals are based on equal-sized groups.

output multiplier changed, and the calculation of each inverse took around 8 minutes.[9] Furthermore, experience shows that it is tough to find intervals to match threshold criteria. This quasi-arbitrary decision adds to the consumption of more time to an approach that already has a characteristic of taking far longer to calculate.

4.2.6. Filter on average transaction amount (AT)

The main advantage of a filter that uses the average transaction amount is that like the filter for holistic accuracy or the Weaver–Thomas Index, its threshold is endogenous to the referent matrix. Thus, the method provides an individual filter value for the base and transformed matrix. The results are similar compared to those from the filter on absolute trade values. The number of trade connections decreases strongly with an increased filter threshold, whereas more than 90% of the total trade values are still represented. There are also differences, however. Due to the structure of EORA26, its average filter value is more moderate than that for the average filter value for EXIOBASE. A comparison of the filter value Θ with the filter method on absolute trade value shows that the results lie between the filter values of 1000 and 500,000 units. A comparison of $10 \times \Theta$ with the filtering on absolute trade values indicates that for EORA26 the results are still between the filter values of 1000 and 500,000 units but for EXIOBASE, the results are comparable with a filter value above 1 million units. The reason for this is that Θ for EORA is much lower compared to that for EXIOBASE and is less sensitive to an increase in the filter threshold.

The filter value applied by Okamoto (2005) (ten times the average transaction amount) would be, depending on the matrix, from 600 to 800 thousand units. Because the filter value is based on trade values, larger trade values are not filtered, as indicated by the low WAD^{rel}.

5. Case study: BSE and foot-and-Mouth disease in the United Kingdom

In this case study, the investigated filter methods are applied to the time series of EXIOBASE from 1995 to 2011 (Stadler et al., 2018). The results of the applied filter methods are not investigated in detail for the whole matrix, as is done previously, but rather for a single sector in a single country, the cattle sector in the United Kingdom (UK). I selected it because outstanding events befell this sector, as can be clearly observed in the data. Both the bovine spongiform encephalopathy (BSE) and the foot-and-mouth disease forced the sector to shut down.

According to the description of the balancing routine 'positive, negative and zero values are maintained throughout the balancing' (Wood et al., 2014, p.151), we are led to the conclusion that only trade connections in the MRIO table exist. They are reported in the base data and no unreported trade connections are inserted via data balancing techniques. That is, I expect that after applying the filter methods, which should consider sector-specific characteristics, the two significant events that shocked this sector should still be observable in the data.

The first case of BSE was reported in the 1980s. It soon started to spread to other animals as well as humans, causing neurological disease (Ainsworth & Carrington, 2000). In March

[9] The calculations are performed on a virtual computer with 192 GB main memory, 2.93 TB hard disk, and an MD EPYC 7402 Server with 2.8 GHz/24-core.

1996, the European Union imposed a worldwide ban on British beef. In 1999, the ban was amended, so that the export of deboned beef and beef products produced under the Date-based Export Scheme (DBES) (European Commission, 2006) was allowed. In order to lift the ban completely, the UK had to fulfill certain conditions, laid out by the European Commission in the TSE[10] Road Map, as a result of which the ban of British beef ended officially in May 2006.

Foot-and-mouth disease was identified as spreading in February 2001. It affects cattle, sheep, pigs, goats, and other ruminants; however, the most serious effects were experienced by the dairy industry. On 21 February 2001, the European Commission imposed a ban on moving any animals susceptible to foot-and-mouth disease and on untreated U.K.-produced products from such livestock. The epidemic lasted 32 weeks; with a last case confirmed on 30 September 2001. A year later in January 2002, the international animal health organization reinstated the U.K. as free of foot-and-mouth disease, and a month later the European Commission lifted meat and animal export restrictions (Eales et al., 2002)

The two mentioned vents can be clearly observed in EXIOBASE data, shown in Figure 1(a) by line Nr.1. This Figure shows the number of sectors to which the cattle sector of the U.K.[11] delivered products. Note that after the ban on beef was amended in 1999, the number of sectors supplied by the U.K.'s cattle sector increased from 232 to 694 sectors. The industry's 2001 collapse due to the foot-and-mouth outbreak can be clearly observed as well, dropping from 532 to 107 sectors. For completeness, the financial crisis is also marked. After applying the filter methods, the clear picture of the events vanishes; compare Figure 1(a). Contrary to expectations, this shows, while filter methods take sector-specific characteristics into consideration, they can smooth unusual deviations. A detailed investigation shows, and as described in the following paragraph, that these results are mostly in line with the previous results.

The trade connections are reduced dramatically, whereas the absolute trade values of the cattle sector do not decrease much. Only with a TLA of 20% does the trade volume decrease by 77% in selected years. For all other filters, the trade volume is above 90%. Also, the European Commission (2006) reported that the amended beef ban in 1999 did not result in any significant exports of beef from the U.K. All the same, it is still worth looking at the results of the other filter methods in detail. Figure 1(a) shows all applied filters in one graph. Figure 1(b) shows only the filters Nr.3 to Nr.11 for a section of the scale. Note in Figure 1(b) that the filter methods of holistic accuracy (Nr.11) and average transaction values (Nr.4) yield similar results; but a comparison with the results of the previous investigation in Section 4 shows that the filter based on average transaction values is much more restricted than is the filter based on holistic accuracy. The filter value of average transaction values varies temporally between 0.208 and 0.244 million EUR, indicating more variation. The fixed filter value of 1.Mil absolute trade values (Nr.3), Leontief's filter of $1/3n$ (Nr.6), and ten times the average transaction value (Nr.5) lead to roughly similar results. Those filter values are also applied in supporting literature (e.g. Aroche-Reyes, 2001; Bierkandt et al., 2014; Okamoto, 2005; Wenz et al., 2014). As already previously indicated, the TLA filter values (Nr.9 and 10), Leontief's filter of $1/n$ (Nr.7) and the filter based on the Weaver–Thomas Index (Nr.8) are the most restrictive.

[10] Transmissible spongiform encephalopathies (TSE) Road Map European Commission, 2005

[11] Note that in the EXIOBASE the country United Kingdom is assigned with the country code GB.

Figure 1. Number of supplied sectors after selected filters by the beef sector in the UK.

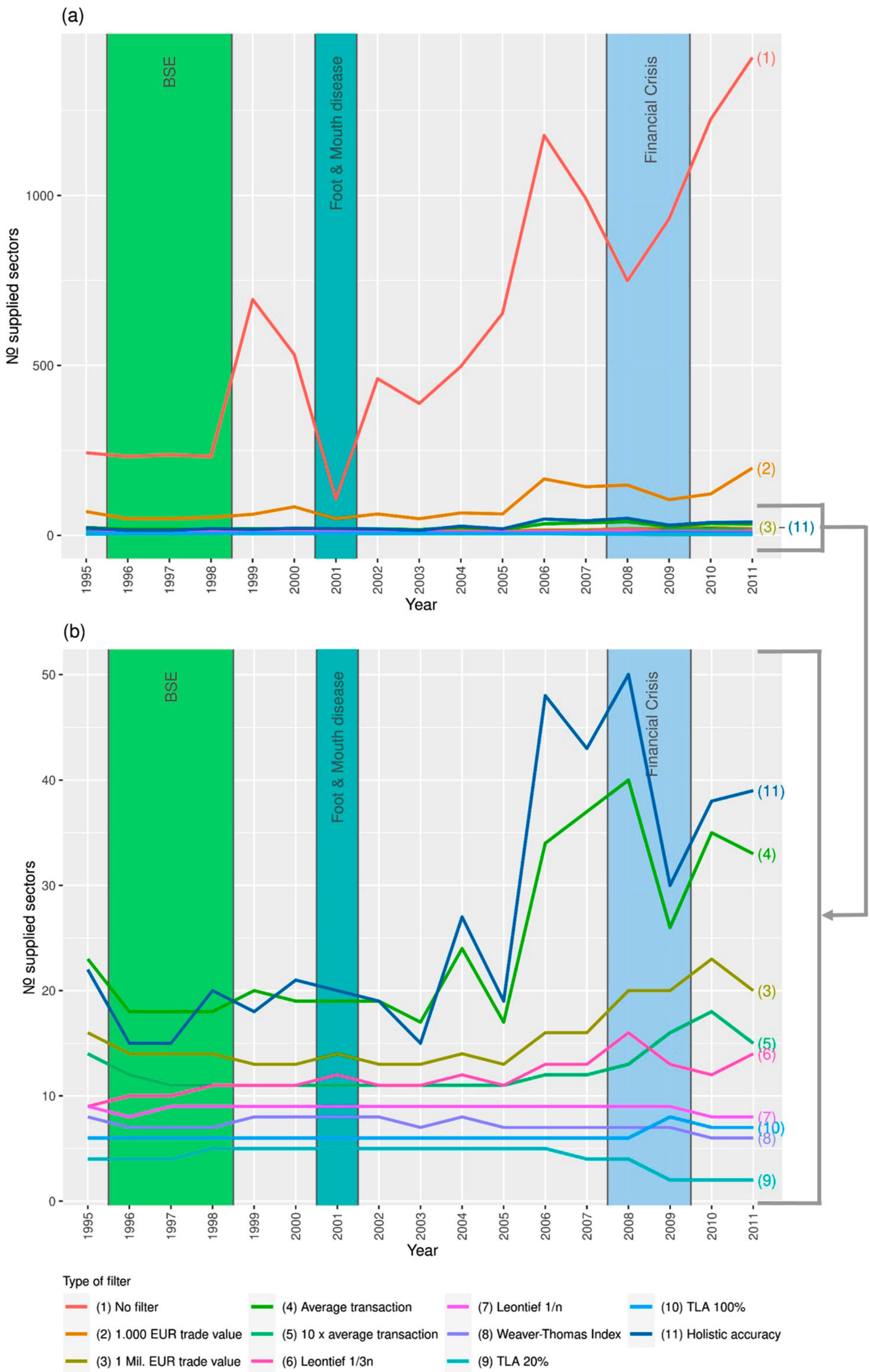

Table 2. Summary of the filter methods.

Filter method	Filter based on:	Consider specific matrix aspects	Strictness of the filter concerning trade connections:	Covering trade values:	Implementation:	Recommendation:
TV	Trade values	No	Adjustable	Very good	Very easy	+
LEO	Coefficients	Partial	High	Medium	Easy	°
WTI	Distribution of the coefficients	Yes	High	Medium	Easy	–
TLA	Impact of coefficients on the output	Yes	High	Medium	Demanding	–
HA	Output multiplier	Yes	Low	Very good	Very demanding	°
AT	Average transaction amount applied to trade values	Yes	Moderate/Adjustable	Very good	Easy	+

Notes: TV, filter on absolute trade value; LEO, Leontief's filter; WTI, Weaver–Thomas Index; TLA, tolerable limits approach; HA, holistic accuracy; AT, filter on average transaction amount. +recommended; °, partially recommended; – , not recommended.

6. Discussion and conclusion

Detailed investigation shows that a large number of small values prevail in MRIO table. For further research, as for example the investigation of the global networks based on MRIO tables, it can be important to answer the question: 'What is an appropriate filter method?'. Prior to the research reported here, no clear answer had yet been provided in the relevant literature. So, I applied different filtering methods to two existing world MRIO tables and compared findings. I extended these results via a case study of the U.K. cattle sector to confirm the broader simulation findings. Table 2 summarizes my investigation into the filtering methods and their characteristics. It also assesses the restrictiveness of the filters in terms of trade connections, how well the trade values are covered, and how easy the method is to implement.

The filter on absolute trade values (TV) is the easiest to implement. It is not matrix specific; that it is not sensitive to such things as number of sectors or countries, total sectoral output, etc. For TV, a filter value of 1 million units yields reasonable results and, hence, seems to be an appropriate filter. Leontief's filter (LEO) is restrictive; only a few trade connections remain after it is applied. Furthermore, by applying it, larger trade values take on widely different values, as indicated by the weighted absolute distance measure. But case study findings suggest relaxing the filters denominator to $1/3n$ can yield reasonable results. This filter also accounts for the number of sectors; but this issue can be misleading as discussed in Section 4.2.2. If more time can be devoted to data preparation, an analysis of the sensitivity of results to different denominators for Leontief's filter is recommended.

The tolerable limits approach (TLA) also takes matrix-specific aspects (e.g. number of sectors or trade volume) into consideration; but the method is designed to assess IO tables for a single economy, so it yields far more restrictive results in an MRIO context. So, TLA is not recommended for MRIO tables since it does not cover interregional or intercountry trade well and its implementation is quite demanding.

A filter based on holistic accuracy (HA) focuses upon output multipliers and also considers many matrix-specific aspects. It is by far the most complicated filter method to implement. It takes a lot of time to find the proper filter value. A comparison with the other filter methods indicates that this filter method is not restrictive and could be preferred to filter out unrealistically small values.

The filter method based on the Weaver–Thomas index (WTI) accounts for the distribution of technology coefficients but is a very restrictive when only a few trade connections exist. The results are similar to those of Leontief's filter with $1/n$. Because the WTI is not variable, the results are similar to LEO. Thus, LEO would be preferred to WTI.

Filtering on the size of the average transaction (AT) has the advantage that for each matrix, an individual filter value is relatively easily derived and is matrix-specific. Its drawback is that it is also very sensitive to the properties of the underlying matrix. In the case study, a filter value of ten times the average transaction yields results similar to those of a filter value of 1 million units and Leontief's filter with a denominator of $1/3n$. It is, therefore, also a preferred filter method, at least for the EXIOBASE. As with Leontief's filter or to a filter of 1 million units, I recommend analyzing the sensitivity of this method to different thresholds to get feel for the data, but as my results and those of Okamoto (2005) show, a filter value of ten times the average transaction tends to work well.

In summary, I do not recommend filtering on the TLA or WTI. Leontief's filter yields reasonable results, but more work should be done on its sensitivity to different denominator values. Of the filter methods compared here, results suggest that those based on total trade values and the average transaction amount should be preferred. The choice between them depends on the priority of matrix-specific characteristics in the analysis being performed.

It is, of course, not possible to find an exact solution. But I was able to compare various solutions and identify 'good' ones. Furthermore, with the broad overview of various filter methods, the reader is now better informed about factors that make filter methods appropriate to a given application. I hope to have raised researchers' awareness of the different filter methods and encourage them to test the sensitivity of results to them, either by altering the thresholds or by applying different filters, in order to assure robust results (e.g. Aroche-Reyes, 2001; Xiao et al., 2017). The purpose of filtering itself could be reduced if MRIO tables were produced without balancing algorithms. This is because network analysis does not require a balanced MRIO table, although IO analysis does. The use of raw values would likely ameliorate the occurrence of spurious small values, so the purpose of filtering would just be that of identifying the most important trade connections. An example would be the global resource accounting model (Wiebe et al., 2012; Wiebe & Lenzen, 2016) where no matrix balancing routine is applied.

Further research should also be conducted to improve upon the set of existing filter methods. For example, the choice of the assumed distribution for the Weaver–Thomas index or a refinement of the tolerable limit approach for MRIO tables. The research indicates, however, that existing and simple methods, such as the filter on absolute trade value or average transaction amount, already yield reasonable results.

Acknowledgments

The author would like to express his appreciation of the time and effort of three anonymous reviewers and the Editor Michael Lahr and acknowledge their very detailed and constructive comments on this paper. Thanks to my supervisor Hans Manner and colleges for the fruitful comments and feedback during the research process.

Disclosure statement

No potential conflict of interest was reported by the author.

Funding

This research was funded by the European Union's Horizon 2020 under grant agreement No. 760891 project CHIC.

References

ABS (2017). Australian bureau of statistics. 5215.0.55.001 – Australian national accounts: Input output tables (Product Details), 2014–15 Explanatory notes.
ABS (2019). Australian bureau of statistics. 5209.0.55.001 – Australian national accounts: Input-output tables 2016–17 Explanatory notes.
Acemoglu D., Carvalho V. M., Ozdaglar A., & Tahbaz-Salehi A. (2012). The network origins of aggregate fluctuations. *Econometrica*, *80*(5), 1977–2016. https://doi.org/10.3982/ECTA9623
Acemoglu D., Ozdaglar A., & Tahbaz-Salehi A. (2016). Networks, shocks, and systemic risk. In Y. Bramoullé, A. Galeotti, B. Rogers (Eds.), *The Oxford handbook of the economics of networks* (pp. 569–607). Oxford University Press.
Ainsworth C., & Carrington D. (2000). BSE disaster: The history. *New Scientist*.
Aroche-Reyes F. (1996). Important coefficients and structural change: A multi-layer approach. *Economic Systems Research*, *8*(3), 235–246. https://doi.org/10.1080/09535319600000017
Aroche-Reyes F. (2001). The question of identifying industrial complexes revisited: A qualitative perspective. In M. L. Lahr, E. Dietzenbacher (Eds.), *Input-output analysis: Frontiers and extensions* (pp. 280–296). Palgrave.
Bierkandt R., Wenz L., Willner S. N., & Levermann A. (2014). Acclimate-a model for economic damage propagation. Part 1: basic formulation of damage transfer within a global supply network and damage conserving dynamics. *Environment Systems and Decisions*, *34*(4), 507–524. https://doi.org/10.1007/s10669-014-9523-4
Campbell J. (1975). Application of graph theoretic analysis to interindustry relationships. *Regional Science and Urban Economics*, *5*(1), 91–106. https://doi.org/10.1016/0166-0462(75)90015-0
Cassetti M. (1995). A new method for the identification of patterns in input-output matrices. *Economic Systems Research*, *7*(4), 363–382. https://doi.org/10.1080/09535319500000002
Chopra S. S., & Khanna V. (2015). Interconnectedness and interdependencies of critical infrastructures in the US economy: Implications for resilience. *Physica A: Statistical Mechanics and Its Applications*, *436*, 865–877. https://doi.org/10.1016/j.physa.2015.05.091
de Mesnard L. (2001). On boolean topological methods of structural analysis. In M. L. Lahr & E. Dietzenbacher (Eds.), *Input-output analysis: Frontiers and extensions* (pp. 268–279). Palgrave.
Dietzenbacher E., van der Linden J. A., & Steenge A. E. (1993). The regional extraction method: EC input-Output comparisons. *Economic Systems Research*, *5*(2), 185–206. https://doi.org/10.1080/09535319300000017
Eales R., Thomas P., Bostock D., Lingard S., Derbyshire I., Burmiston A., & Kitson H. (2002). The 2001 outbreak of foot and mouth disease (Technical report). National Audit Office.
European Commission (2005). *The TSE roadmap*. Brussels.
European Commission (2006). *BSE: UK beef embargo to be lifted*. Brussels. IP/06/278.
Gaulier G., & Zignago S. (2010). BACI: international trade database at the product-Level (the 1994-2007 version). *SSRN Electronic Journal*.
Holub H. W., Schnabl H., & Tappeiner G. (1985). Qualitative input-Output analysis with variable filter. *Zeitschrift für die gesamte Staatswissenschaft/Journal of Institutional and Theoretical Economics*, *141*, 282–300. https://www.jstor.org/stable/40750836
Jensen R. C. (1980). The concept of accuracy in regional input-Output models. *International Regional Science Review*, *5*(2), 139–154. https://doi.org/10.1177/016001768000500203

Jensen R. C., & West G. R. (1980). The effect of relative coefficient size on input-Output multipliers. *Environment and Planning A: Economy and Space*, *12*(6), 659–670. https://doi.org/10.1068/a120659
Jilek J. (1971). The selection of the most important coefficients. *Economic Bulletin for Europe*, *23*, 86–105.
Lahr M. L. (2001). Appendix 3 some measures for comparing input-output matrices. In M. L. Lahr, E. Dietzenbacher (Eds.), *Input-output analysis: Frontiers and extensions* (pp. 236–242). Palgrave.
Lenzen M. (2019). Aggregating input–output systems with minimum error. *Economic Systems Research*, *31*(4), 594–616. https://doi.org/10.1080/09535314.2019.1609911
Lenzen M., Kanemoto K., Moran D., & Geschke A. (2012). Mapping the structure of the world economy. *Environmental Science & Technology*, *46*(15), 8374–8381. https://doi.org/10.1021/es300171x
Lenzen M., Moran D., Kanemoto K., & Geschke A. (2013). Building EORA: A global multi-Region input-Output database at high country and sector resolution. *Economic Systems Research*, *25*(1), 20–49. https://doi.org/10.1080/09535314.2013.769938
Leontief W. W. (1965). The structure of the U.S. economy. *Scientific American*, *212*(4), 25–35. https://doi.org/10.1038/scientificamerican0465-25
Luu D. T., Napoletano M., Fagiolo G., Roventini A., & Sgrignoli P. (2017). Uncovering the network complexity in input-output linkages among sectors in European Countries. *ISIGrowth Working Paper* (pp. 1–27).
Miller R. E. (1966). Interregional feedback effects in input-output models: Some preliminary results. *Papers of the Regional Science Association*, *17*(1), 105–125. https://doi.org/10.1007/BF01982512
Miller R. E. (1969). Interregional feedback effects in input-output models: Some experimental results. *Economic Inquiry*, *7*(1), 41–50. https://doi.org/10.1111/ecin.1969.7.issue-1
Miller R. E., & Blair P. (1983). Estimating state-level input-output relationships from U.S. multiregional data. *International Regional Science Review*, *8*(3), 233–254. https://doi.org/10.1177/016001768300800304
Miller R. E., & Blair P. D. (2009). *Input-Output analysis: Foundations and extensions*. Cambridge University Press.
Miller R. E., & Lahr M. L. (2001). *A taxonomy of extractions*. In M. L. Lahr, R. E. Miller (Eds.), *Regional science perspectives in economic analysis: A festschrift in memory of Benjamin H. Stevens* (pp. 407–441). Elsevier Science.
Okamoto N. (2005). Agglomeration, intraregional and interregional linkages in China. In N. Okamoto & T. Ihara (Eds.), *Spatial structure and regional development in China: An interregional input-output approach* (pp. 128–153). Palgrave Macmillan UK.
Paelinck J., De Caevel J., & Degueldre J. (1965). Analyse quantitative de certaines phénomenes du développment régional polarisé: Essai de simulation statique d'itérarires de propogation. *Bibliothèque de l'Institut de Science économique*, *7*, 341–387.
Richardson H. W. (1972). *Input-output and regional economics*. Wiley.
Sawyer C. H., & Miller R. E. (1983). Experiments in regionalization of a national input-output table. *Environment and Planning A: Economy and Space*, *15*(11), 1501–1520. https://doi.org/10.1068/a151501
Schintke J., & Stäglin R. (1988). Important input coefficients in market transaction tables and production flow tables. In M. Ciaschini (Ed.), *Input-output analysis current developments* (pp. 43–60). Chapman and Hall.
Schnabl H. (1994). The evolution of production structures, analyzed by a multi-layer procedure. *Economic Systems Research*, *6*(1), 51–68. https://doi.org/10.1080/09535319400000004
Schultz S. (1976). Intersectoral comparisons as an approach to the identification of key sectors. In K. R. Polenske & J. V. Skolka (Eds.), *Advances in input-output analysis* (pp. 137–159). Ballinger Publishing Company.
Songling X., & Gould P. (1991). The grad field of input-output models and the nature of coefficients. *Economic Systems Research*, *3*(4), 367–378. https://doi.org/10.1080/09535319100000026
Stadler K., Wood R., Bulavskaya T., Södersten C. J., Simas M., Schmidt S., Usubiaga A., Acosta-Fernández J., Kuenen J., Bruckner M., Giljum S., Lutter S., Merciai S., Schmidt J. H., Theurl M. C., Plutzar C., Kastner T., Eisenmenger N., Erb K. H., & Tukker A. (2018). EXIOBASE 3: developing a

time series of detailed environmentally extended multi-Regional input-Output tables: EXIOBASE 3. *Journal of Industrial Ecology*, *22*(3), 502–515. https://doi.org/10.1111/jiec.2018.22.issue-3
Strassert G. (1968). Zur Bestimmung strategischer Sektoren mit Hilfe von Input-Output Modelle. In *Jahrbücher Für Nationalökonomie und Statistik* (Vol. 182. pp. 211–215). Gustav Fischer Verlag Stuttgart.
Szyrmer J. M. (1984). Estimating interregional input-output tables with RAS. *Working paper*, Social Science Data Center, University of Pennsylvania, unpublished.
Tarancón M., Callejas F., Dietzenbacher E., & Lahr M. L. (2008). A revision of the tolerable limits approach: searching for the important coefficients. *Economic Systems Research*, *20*(1), 75–95. https://doi.org/10.1080/09535310801890714
Thakur S. K. (2011). Fundamental economic structure and structural change in regional economies: A methodological approach. *Région Et Développement*, *33*, 9–38.
Timmer M. P., Dietzenbacher E., Los B., Stehrer R., & de Vries G. J. (2015). An illustrated user guide to the world input–output database: the case of global automotive production. *Review of International Economics*, *23*(3), 575–605. https://doi.org/10.1111/roie.2015.23.issue-3
Tumminello M., Micciche S., Lillo F., Piilo J., & Mantegna R. N. (2011). Statistically validated networks in bipartite complex systems. *PloS One*, *6*, 1–11. https://doi.org/10.1371/journal.pone.0017994
Wenz L., Willner S. N., Bierkandt R., & Levermann A. (2014). Acclimate- a model for economic damage propagation. Part II: A dynamic formulation of the backward effects of disaster-induced production failures in the global supply network. *Environment Systems and Decisions*, *34*(4), 525–539. https://doi.org/10.1007/s10669-014-9521-6
Wenz L., Willner S. N., Radebach A., Bierkandt R., Steckel J. C., & Levermann A. (2015). Regional and sectoral disaggregation of multi-regional input-Output tables – A flexible algorithm. *Economic Systems Research*, *27*(2), 194–212. https://doi.org/10.1080/09535314.2014.987731
Wiebe K. S., Bruckner M., Giljum S., & Lutz C. (2012). Calculating energy-related CO2 emissions embodied in international trade using a global input–output model. *Economic Systems Research*, *24*(2), 113–139. https://doi.org/10.1080/09535314.2011.643293
Wiebe K. S., & Lenzen M. (2016). To RAS or not to RAS? What is the difference in outcomes in multi-regional input–output models?. *Economic Systems Research*, *28*(3), 383–402. https://doi.org/10.1080/09535314.2016.1192528
Wood R., Stadler K., Bulavskaya T., Lutter S., Giljum S., A. de Koning, Kuenen J., Schütz H., Acosta-Fernández J., Usubiaga A., Simas M., Ivanova O., Weinzettel J., Schmidt J., Merciai S., & Tukker A. (2014). Global sustainability accounting-developing EXIOBASE for multi-Regional footprint analysis. *Sustainability*, *7*(1), 138–163. https://doi.org/10.3390/su7010138
Xiao H., Sun T., Meng B., & Cheng L. (2017). Complex network analysis for characterizing global value chains in equipment manufacturing. *PloS One*, *12*(1), e0169549. https://doi.org/10.1371/journal.pone.0169549
Xiao WW., Wang Ll., Zhang Zy., & Wang Cw. (2019). Identify and analyze key industries and basic economic structures using interregional industry network. *Cluster Computing*, *22*(S4), 10041–10051. https://doi.org/10.1007/s10586-017-1067-1
Xu X. Z., & Liu S. D. (2013). The method of selecting the key freight transport Industries based on weaver-Thomas index. *Advanced Materials Research*, *779-780*, 981–984. https://doi.org/10.4028/www.scientific.net/AMR.779-780
Zhang Z., Chen X., Xiao W., & Qi G. (2016). Identifying and analyzing strong components of an Industrial network based on cycle degree. *Scientific Programming*, *2016*, 1–11. http://dx.doi.org/10.1155/2016/7340859