

Explaining value chain differences in MRIO databases through structural path decomposition

Anne Owen, Richard Wood, John Barrett & Andrew Evans

To cite this article: Anne Owen, Richard Wood, John Barrett & Andrew Evans (2016) Explaining value chain differences in MRIO databases through structural path decomposition, *Economic Systems Research*, 28:2, 243-272, DOI: [10.1080/09535314.2015.1135309](https://doi.org/10.1080/09535314.2015.1135309)

To link to this article: <https://doi.org/10.1080/09535314.2015.1135309>



© 2016 The Author(s). Published by Taylor & Francis.



[View supplementary material](#)



Published online: 06 Feb 2016.



[Submit your article to this journal](#)



Article views: 2729



[View related articles](#)



[View Crossmark data](#)



Citing articles: 30 [View citing articles](#)

Explaining value chain differences in MRIO databases through structural path decomposition

Anne Owen^a, Richard Wood^b, John Barrett^a and Andrew Evans^c

^aSchool of Earth and Environment, Sustainability Research Institute, University of Leeds, Leeds, UK;

^bIndustrial Ecology Programme, Norwegian University of Science and Technology, Trondheim, Norway;

^cSchool of Geography, University of Leeds, Leeds, UK

ABSTRACT

Many multiregional input–output (MRIO) databases are used to calculate consumption-based accounts. Results feature in climate policy discussion on emissions reduction responsibilities; yet studies show that outcomes produced by each database differ. This paper compares the emissions associated with value chains from Eora, EXIOBASE, GTAP and WIOD. Structural path analysis identifies the largest paths in each database and the differences in common paths are calculated. For the top 100 value chain paths that contain the largest difference, structural path decomposition is used to identify the contribution each part of the value chain makes towards the difference. The results identify and quantify key flows that are the cause of difference in the databases. From these, we can conclude that key MRIO database construction decisions, such as using the residence or territorial principle for emissions allocation and whether energy spends are reallocated based on physical data, are the major causes of differences.

ARTICLE HISTORY

Received 29 July 2015

Accepted 19 December 2015

KEYWORDS

Multiregional input–output databases; variation; global value chains; structural path decomposition; consumption-based accounting; model comparison

1. Introduction

Understanding how to develop policy to mitigate greenhouse gas (GHG) emissions has become difficult in an increasingly globalised world. To appreciate the role of trade in terms of emissions, calculations involving multiregional input–output (MRIO) databases have become the dominant and most progressive method. These databases centre on the evaluation and manipulation of trade flows between regions and industrial sectors, using a flow matrix approach. For example, the flow of steel from steel production into car manufacturing is associated with the CO₂ consequent upon that use, allowing the full supply chain emissions of cars to be calculated. The number and types of policy applications of MRIO suggested by both academics and policy makers is growing exponentially. There are a number of leading databases available and each database produces a different set of consumption-based accounts (CBA). Each country's difference from the multi-model mean CBA is shown in Figure 1. However, there has been little appreciation as to why they produce different results. Owen et al. (2014) used structural decomposition techniques to

CONTACT Anne Owen  a.owen@leeds.ac.uk

© 2016 The Author(s). Published by Taylor & Francis.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

the emission estimates of groups of goods and services is paramount. One way to understand the reliability is to focus on the consistency of supply chain impacts across the models. Hence the first aim of this paper is to use structural path analysis (SPA) to find, for each database pairing, the paired value chains that exhibit the largest differences. For example, the value chain that describes the emissions associated with the electricity used to make steel that ends up in cars bought by German consumers might not be the largest path in calculating the CBA for Germany using Eora or WIOD. However, when the size of this particular path is compared between the two database calculations, it might have a large difference. Once the one hundred largest path differences (PD) are calculated for every common country, for each database pairing, the second aim is to use decomposition techniques to determine which part of the value chain is responsible for the highest portion of the difference. This technique has become known as structural path decomposition (SPD) (Wood and Lenzen, 2009).

1.1. Structural path analysis

SPA is a technique that decomposes a consumption-based account to the sum of an infinite number of production chains – sometimes called paths. Wood and Lenzen, (2003, p. 371) describe this process ‘unravelling the Leontief inverse using its series expansion’. The SPA technique was first described by Defourny and Thorbeck (1984) and Crama et al. (1984). SPA can be used to find those production chains that contribute most to a particular CBA. Paths are categorised according to their length. For example, a zero-th order path represents an industry’s direct on-site emissions arising from final demand of the product produced by that particular industry. This could be the emissions from transport in providing a transport service. A first order path has one further step in the chain: for example the emissions from steel production that are used to make cars for final demand. Most SPA studies rank these chains in order of their importance. Because there are an infinite number of paths of decreasing importance that sum to the total CBA, most authors will display the top 20 or so chains. Writing in 2006, Peters and Hertwich state that there are very few I–O studies that apply SPA and that hybrid Life Cycle Assessment techniques are a more popular method employed to consider production chains. By 2016, this is still the case – SPA methods remain relatively underexposed, particularly amongst research using MRIO databases.

Wood and Lenzen (2003) use SPA and a 1995 I–O database for Australia to compare the CBA of two Australian research institutions. Their analysis reveals a large proportion of the two institutions’ ecological footprint impacts occurring upstream in first or second order paths. Using the same database, Lenzen (2003) furthers this work to analyse the Australian economy as a whole and considers CBAs calculated using energy, land, water, GHG, NO_x and SO₂ emissions. Lenzen (2003) demonstrates that when considering energy and emissions rather than land use, the zero-th order paths dominate the rankings. The reason for this is that direct land use only applies to a few industrial sectors. A production chain has to start with one of these sectors to show as having significant impact. This means that chains will often have to be least first order to link to the land using sectors. There is significant direct emissions use for a wider proportion of industrial sectors meaning that many zero-th order paths will be significant. The advantage of a emissions-based study is that the largest paths will be relatively short and quick to find during the SPA procedure. Both Lenzen’s (2003) and Peters and Hertwich’s (2006) analyses, of Australia and Norway, respectively,

find that zero-th order paths involving electricity, metals, chemicals and transport services are significant.

SPA can also be used to consider 'sustainable chain management'. Foran et al. (2006) explain that by identifying the all processes in the production of a product, it is possible to decide whether to focus on improving direct onsite production processes or the indirect contributions further along the supply chain. Foran et al.'s (2006) study was based on a request by the Australian government to look at the Triple-Bottom-Line performance of industry sectors.

Treloar (1997) uses SPA to investigate the embodied energy paths of the Australian residential buildings sector. Based on his investigation, Treloar (1997) concludes that SPA techniques can improve the completeness of process-based analysis of product impacts and suggests such a hybrid technique. However, taking process-based life-cycle and combining this with I–O data to account for the missing information will result in the double-counting of certain paths. The identification and correction of double-counted paths has been the subject of much debate between Strømman et al. (2009), Lenzen and Crawford (2009) and Strømman (2009). Acquaye et al. (2011) use this hybrid methodology, specifically combining a UK focused two-region MRIO database with process LCA data, to consider the upstream paths that contribute to the production of biofuels. The authors discuss how SPA has been used in this case to identify carbon 'hot spots', or rather the highest carbon intensity path of the upstream supply chain for biodiesel.

Other methodological enhancements include work by Sonis et al. (1997, p. 278) whose block SPA technique identifies 'paths of influence' and can reveal the 'finer structure of economies' using a macro level approach rather than the micro level of each individual path. And, beyond the field of CBA, SPA has also been suggested by Suh (2005) and then implemented by Lenzen (2007) to explore and understand relationships between components of an ecosystem. Suh (2005, p. 256), however, warns that applying ideas from the field of economics to ecology might not always be appropriate but notes that the I–O system is an 'efficient way of presenting . . . data for a network structure'.

1.2. Structural path decomposition

SPD was developed by Wood and Lenzen (2009) as a combination of decomposition analyses and SPA. Wood and Lenzen (2009) use SPD to understand changes in a production chain between two points in time. Where decomposition analyses assign proportions of the difference in CBA to elements in the environmentally extended Leontief input–output equation, SPD assigns difference proportions to elements in a product's supply chain. For example, the largest difference in a chain between times t_0 and t_1 could occur in a zero-th order path such as the onsite electricity emissions making an electricity final demand product or a first order path, such as the emissions from livestock that are used to make food products for final demand. In addition to identifying the chains that contribute most to the difference, SPD identifies which part of the chain has the highest difference associated with it. For example, in the second order path representing the livestock emissions associated with final demand for food, the difference between this path in t_0 and t_1 can be shared between the three parts of the chain: the emissions intensity of livestock production; the amount of livestock needed to make a food product; and the amount of food product bought by final demand consumers.

Wood and Lenzen (2009) use the logarithmic mean divisia index (LMDI) decomposition technique for the SPD methodology and apply it to Australian I–O tables for 1995 and 2005. Wood and Lenzen (2009) find that between 1995 and 2005, the largest changes in emissions production paths involved livestock and electricity. The element most responsible for difference tends to be either a change in level of domestic final demand or a change in level of demand for export.

Since Wood and Lenzen's (2009) initial paper, there have been very few applications of the technique in the literature. Oshita (2012) uses SPD to look at changes in CO₂ emissions in Japanese supply chains between 1990 and 2000 and Gui et al. (2014) consider changes in CO₂ emissions in Chinese supply chains between 1992 and 2007. Both examples use SPD to explain a change in emissions over time but rather than use the LMDI technique, both Oshita (2012) and Gui et al. (2014) opt for polar decompositions.

Clearly, there is an opportunity for SPD techniques to be applied to different MRIO systems rather than different time frames. The work presented in this paper may present the first application of SPD for this use. In addition there is also an option to explore using the Dietzenbacher and Los (D&L) (Dietzenbacher and Los, 1998) or Shapely–Sun (S–S) (Sun, 1998) decomposition technique within the SPD calculations, which are considered more suitable than polar decompositions (de Boer, 2009).

2. Data and methods

As explained in the paper by Owen et al. (2014), in order to compare Eora, EXIOBASE, GTAP and WIOD, the databases need to be the same size and have the same structure. These means that each of the databases must be mapped on to a common classification (CC) containing a common set of regions and sectors, presented in the same order, format and currency. This paper adopts the same CC system, of 41 regions and 17 sectors, used in Owen et al. (2014) and Steen-Olsen et al. (2014) and the reader is directed to these papers for more information as to how the aggregation system was generated. This study uses a CC that is slightly different to those described in Owen et al. (2014) and Steen-Olsen et al. (2014) in that a symmetric input–output table (SIOT) structure is preferred to the supply and use table (SUT) format. Finding structural paths with MRIO databases in SUTs formats adds additional complexity to the purpose of uncovering database differences so it was decided to convert the aggregated databases into an industry-by-industry SIOT format, using the fixed product sales structure assumption for use in this paper. This paper chooses to use data from the year 2007 since this is the one year common to all four databases. In their comparison of MRIO databases, Moran and Wood (2014) use a harmonised CO₂ emissions vector. For this study we take the CO₂ emissions data provided by each database supplier to allow for investigation into whether the emissions or economic data is more responsible for difference in value chains. We calculate the largest paths in the following databases:

- Eora CC – Eora mapped to the common classification
- EXIOBASE CC – EXIOBASE mapped to the common classification
- GTAP CC – GTAP mapped to the common classification
- WIOD CC – WIOD mapped to the common classification

We then compare the same paths in the corresponding databases (based on the same aggregations) to find the top 100 paths for each country with the largest path difference.

2.1. SPD equations used

A series expansion is used to calculate the largest paths in each database:

$$\begin{aligned} \mathbf{q} = & \sum_{i=1}^n e_i y_i + \sum_{i=1}^n e_i \sum_{j=1}^n a_{ij} y_j + \sum_{i=1}^n e_i \sum_{k=1}^n a_{ik} \sum_{j=1}^n a_{kj} y_j \\ & + \sum_{i=1}^n e_i \sum_{l=1}^n a_{il} \sum_{k=1}^n a_{lk} \sum_{j=1}^n a_{kj} y_j + \dots \end{aligned} \quad (1)$$

adapted from Wood and Lenzen (2003). \mathbf{q} is the total consumption-based emissions, \mathbf{e} is the emissions intensity vector, \mathbf{A} is the direct requirements matrix and \mathbf{y} is the vector of total final demand for the specific country. i , j , k and l are component sectors. A first order path from sector i into sector j is calculated by $e_i a_{ij} y_j$. A second order path from sector i via sector k into sector j is calculated by $e_i a_{ik} a_{kj} y_j$ and so on (Peters and Hertwich, 2006).

For the SPD, rather than find the path difference associated with the elements from \mathbf{e} and \mathbf{A} , it was thought more useful to consider that \mathbf{e} is constructed from the emissions vector \mathbf{f} divided by total output \mathbf{x} and that each element of \mathbf{A} , a_{ij} , is the corresponding element of the transactions element \mathbf{Z} , z_{ij} divided by the corresponding column sum, or rather total output element x_j .

This means that zero-th, first, second and third value chains can be characterised thus:

$$q_{0\text{th}} = f_i \cdot x_i^{-1} \cdot y_i, \quad (2)$$

$$q_{1\text{st}} = f_i \cdot x_i^{-1} \cdot z_{ij} \cdot x_j^{-1} \cdot y_j, \quad (3)$$

$$q_{2\text{nd}} = f_i \cdot x_i^{-1} \cdot z_{ij} \cdot x_j^{-1} \cdot z_{jk} \cdot x_k^{-1} \cdot y_k, \quad (4)$$

$$q_{3\text{rd}} = f_i \cdot x_i^{-1} \cdot z_{ij} \cdot x_j^{-1} \cdot z_{jk} \cdot x_k^{-1} \cdot z_{kl} \cdot x_l^{-1} \cdot y_l. \quad (5)$$

The difference can now be interpreted to consider the effect that the emissions vector has on its own rather than being combined with the effect of total output. However, we must be aware that decompositions will sometimes show very large positive effects for \mathbf{f} and negative effects for \mathbf{x}^{-1} which, in effect, cancel each other. In these cases, we must ensure to consider the effect of the other factors, even though their net effects are smaller. In addition it is also easier to interpret the difference between individual elements in \mathbf{Z} rather than in \mathbf{A} where they intrinsically linked to the remainder of the items in the column because each item shows the proportion of the column sum.

Dietzenbacher and Los (2000) warn that decomposition analyses need to be treated with care due to the dependency problem. A decomposition equation assumes that each term is independent of each other term. However, the authors point out in their example that ‘changes in intermediate input coefficient and in value added coefficient affect each other’ (Dietzenbacher and Los, 2000, p. 4). Decomposition of consumption-based emissions often requires the calculation of the emissions per unit of output. It is not appropriate to assume that a change in emissions efficiency can occur independently of the technology

matrix used to calculate the Leontief inverse. A solution to the dependency problem is suggested by Dietzenbacher and Los (2000) but most decomposition studies do not address it. In fact, few, with the exception of Hoekstra and van der Bergh (2002) and Minx et al. (2011), mention the issue. The equation presented above splits emissions efficiency into the component parts \mathbf{f} and \mathbf{x}^{-1} , this removes the efficiency vector from the equation. This amendment does not follow the proposed form suggested by Dietzenbacher and Los (2000) for cases with dependent determinants. And by introducing \mathbf{Z} and \mathbf{x}^{-1} as a substitute for \mathbf{A} , the dependency issue remains since \mathbf{x} is directly dependent on \mathbf{Z} . There is no simple way of amending the terms to create independency and we highlight that the dependency issue is problematic for all decompositions that assess changes in emissions and energy (Minx et al., 2011).

Splitting \mathbf{e} into the separate elements \mathbf{f} and \mathbf{x}^{-1} , and splitting \mathbf{A} into \mathbf{Z} and \mathbf{x}^{-1} means that paths of zero-th order now contain three elements rather than two. Fourth order paths, which can still give large emissions values, now contain eleven elements rather than six. The D&L (Dietzenbacher and Los, 1998) structural decomposition approach, used in Owen et al. (2014) is too complex for an eleven element comparison. The S-S (Sun, 1998) approach, is instead used to decompose the difference in paths to each element in the value chain equation. S-S is equivalent to the mean effect calculated by D&L but it does not provide the full range of equivalent decompositions. This means that we cannot comment on the variation associated with the contributitional effect to the difference for each term. The general format for PD for paths of zero-th to third order value chains is shown in Equations 6–9, respectively.

$$PD_{0th} = f_{effect} + x_{effect}^{-1} + y_{effect}, \tag{6}$$

$$PD_{1st} = f_{effect} + x_{effect}^{-1} + Z_{effect} + x_{effect}^{-1} + y_{effect}, \tag{7}$$

$$PD_{2nd} = f_{effect} + x_{effect}^{-1} + Z_{effect} + x_{effect}^{-1} + Z_{effect} + x_{effect}^{-1} + y_{effect}, \tag{8}$$

$$PD_{3rd} = f_{effect} + x_{effect}^{-1} + Z_{effect} + x_{effect}^{-1} + Z_{effect} + x_{effect}^{-1} + Z_{effect} + x_{effect}^{-1} + y_{effect}. \tag{9}$$

For the general case $x = y_1 y_2 \dots y_n$, the general format for the S-S decomposition equation is:

$$y_{ieffect} = \frac{x_0}{y_{0i}} \Delta y_i + \sum_{j \neq i} \frac{x_0}{2y_{0,i}y_{0,j}} \Delta y_i \Delta y_j + \sum_{j \neq i \neq k} \frac{x_0}{3y_{0,i}y_{0,j}y_{0,k}} \Delta y_i \Delta y_j \Delta y_k + \dots + \frac{1}{n} \Delta y_1 \Delta y_2 \dots \Delta y_n. \tag{10}$$

3. Results

3.1. Structural path analysis

To illustrate the results produced by a SPA we first use the example of the UK value chains from the GTAP CC and WIOD CC databases. Table 1 shows the top 20 value chains from GTAP.

The largest path in the aggregated GTAP databases for the UK is the path representing the emissions from GBR electricity, gas and water supply that go directly to the final

Table 1. Top 20 largest paths from the GTAP CC for the UK.

Rank	MtCO ₂	Order	Sector 1	Sector 2	%
1	69.9	0	GBR Elec, Gas and Water		11.1
2	36.1	0	GBR transport		5.7
3	18.9	1	GBR Elec, Gas and Water	GBR PDEH ^a	3.1
4	12.5	1	GBR Elec, Gas and Water	GBR Transport	2.0
5	6.4	1	GBR Transport	GBR PDEH	1.0
6	5.8	1	GBR Transport	GBR Transport	0.9
7	5.4	1	GBR Elec, Gas and Water	GBR Food and Drink	0.9
8	5.3	1	GBR Elec, Gas and Water	GBR Elec, Gas and Water	0.8
9	5.2	0	GBR Food and Drink		0.8
10	5.1	0	GBR PDEH		0.8
11	4.8	0	GBR Coke and Petrol		0.8
12	4.4	1	GBR Elec, Gas and Water	GBR Business Services	0.7
13	3.9	0	ROW Transport		0.6
14	3.7	0	GBR Business Services		0.6
15	3.6	1	GBR Transport	GBR Business Services	0.6
16	3.3	1	GBR Transport	GBR Food and Drink	0.5
17	3.2	1	GBR Elec, Gas and Water	GBR Trade	0.5
18	3.0	1	GBR Elec, Gas and Water	GBR Coke and Petrol	0.5
19	2.5	0	USA Transport		0.4
20	2.4	0	ROW Coke and Petrol		0.4
Rest	369.5				58.2

^aPublic Administration, Defence, Education and Health.

demand for that product. This path represents 11.1% of the total CBA for the UK. All of the paths in the top 20 are either zero-th or first order paths. This fits with the findings of Lenzen (2003) who suggests that for SPA using energy and emissions data, most of the large paths are zero-th and first order. The top 20 represents 42% of the overall footprint. Paths originating from the electricity, gas and water supply industry contribute to significant portion of the largest paths. These sectors also featured highly in Peters and Hertwich's (2006) SPA of Norway.

For the corresponding WIOD data, shown in Table 2, the two largest paths are the same but the path in third place is the eighth largest in the GTAP system. Similarly the third largest path in the GTAP data is eighth largest for WIOD. The next stage is to find the largest differences between corresponding paths. For example, the difference between the zero-th order path from the GBR electricity, gas and water supply in the GTAP and WIOD systems is 2.4 MtCO₂. This path is the largest in both tables, but the *difference* may not be the largest. To find the largest differences, we need to look beyond the top 20 paths. To identify the top 100 PD the top 1000 zero-th, first, second, third and fourth order paths were found using GTAP and WIOD. Matching path descriptions were found for each order and the difference calculated. PD were then ranked and any outside the top 100 discarded.

The CBA for the UK as calculated by the GTAP and WIOD databases using the CC system differs by 27.6 MtCO₂, with WIOD calculating the footprint to be higher. Table 3 shows the top 20 value chain differences. The largest difference is for the path from the emissions associated with GBR Transport to the final demand for that same sector. For this path, GTAP calculates higher than WIOD. The second largest path is that from the emissions associated with the GBR electricity, gas and water supply industry that are used for intermediate demand for the same sector and final demand for the same sector. This path is 14.3 MtCO₂ larger in WIOD than in GTAP (hence negative in the Δ MtCO₂ column). Because this path is a first order path, it contains an interaction with a cell in the **Z** matrix.

Table 2. Top 20 largest paths from the WIOD CC for the UK.

Rank	MtCO ₂	Order	Sector 1	Sector 2	Sector 3	%
1	72.3	0	GBR Elec, Gas, Water			10.9
2	20.9	0	GBR Transport			3.2
3	19.7	1	GBR Elec, Gas, Water	GBR Elec, Gas, Water		3.0
4	15.5	0	GBR PDEH			2.3
5	13.0	0	ROW Coke & Petrol			2.0
6	11.2	1	GBR Transport	GBR Trade		1.7
7	9.7	0	GBR Trade			1.5
8	9.5	1	GBR Elec, Gas, Water	GBR PDEH		1.4
9	6.6	0	GBR Coke & Petrol			1.0
10	6.4	0	ROW Construction			1.0
11	5.7	0	GBR Metals			0.9
12	5.4	2	GBR Elec, Gas, Water	GBR Elec, Gas, Water	GBR Elec, Gas, Water	0.8
13	4.7	1	GBR Coke & Petrol	GBR PDEH		0.7
14	4.5	1	GBR Transport	GBR PDEH		0.7
15	4.4	1	GBR Elec, Gas, Water	GBR Trade		0.7
16	3.9	0	ROW Manufacturing			0.6
17	3.5	1	GBR Transport	GBR Transport		0.5
18	3.4	0	GBR Food & Drink			0.5
19	2.6	2	GBR Elec, Gas, Water	GBR Elec, Gas, Water	GBR PDEH	0.4
20	2.5	0	GBR Agriculture			0.4
Rest	375.9					56.7

Table 3. Top 20 PD for the UK from GTAP CC and WIOD CC.

Rank	ΔMtCO ₂	Order	Sector 1	Sector 2	Sector 3
1	15.3	0	GBR Transport		
2	-14.3	1	GBR Elec, Gas, Water	GBR Elec, Gas, Water	
3	-10.6	0	ROW Coke & Petrol		
4	-10.3	0	GBR PDEH		
5	10.1	1	GBR Elec, Gas, Water	GBR PDEH	
6	-8.3	0	GBR Trade		
7	-4.9	2	GBR Elec, Gas, Water	GBR Elec, Gas, Water	GBR Elec, Gas, Water
8	-4.7	0	GBR Construction		
9	-4.4	0	GBR Metals		
10	4.0	1	GBR Elec, Gas, Water	GBR Food and Drink	
11	-4.0	1	ROW Coke and Petrol	GBR PDEH	
12	3.6	0	ROW Transport		
13	-3.4	0	ROW Manufacturing		
14	2.7	1	GBR Transport	GBR Food and Drink	
15	2.4	1	GBR Elec, Gas, Water	GBR Business	
16	-2.4	0	GBR Elec, Gas, Water		
17	2.3	1	GBR Transport	GBR Transport	
18	-2.3	0	GBR Agriculture		
19	-2.3	1	GBR Mining	GBR Elec, Gas, Water	
20	2.0	1	GBR Elec, Gas, Water	GBR Coke and Petrol	

In addition, since this path difference is larger than the path difference associated with the zero-th order path from GBR electricity, gas and water supply (ranked number 16 in Table 3), one would assume that it is data from the transactions matrix causing difference. The total 27.6 MtCO₂ difference between GTAP and WIOD is the sum of thousands of PD, both positive and negative.

The next stage is to find out which element in the series expansion equation, used to calculate the size of a value chain, is responsible for the majority of the difference in paths and to calculate the percentage contribution each element makes to the overall difference.

3.2. Structural path decomposition

Table 4 shows the elements in the emissions vector \mathbf{f} , the inverse output vector \mathbf{x}^{-1} , the transactions matrix \mathbf{Z} and the final demand vector \mathbf{y} from GTAP and WIOD that make up the paths shown in Table 3. As Table 3 shows, the path with the largest difference between GTAP and WIOD is the value chain of emissions for transport that go to make a final demand of the same product. Table 4 shows that the industrial emissions associated with the UK transport sector are 131.4 MtCO₂ in GTAP and 91.9 MtCO₂ in WIOD. The inverse output values are 3.35×10^{-6} and 4.27×10^{-6} . Final demand of UK transport by UK consumers is 82,228 million US dollars (USD) in GTAP and 53,326 USD in WIOD. Each of the \mathbf{f} , \mathbf{x}^{-1} and \mathbf{y} elements differ between the two databases, but which element contributes the most to the path difference of 15.3 MtCO₂? From Table 3, we suspected that the second largest path difference might have something to do with the differences in the \mathbf{Z} matrix. Table 4 shows that there is a large difference between the value from GTAP's \mathbf{Z} matrix (6994) and the value from WIOD's (46,163). But again, how much of the overall difference can be explain by this? SPD is used to calculate the contribution each element in the path makes towards this difference and the results are shown in Table 5.

As expected, the second row of Table 5 reveals that the element that contributes most to the path difference is the element in the transactions matrix \mathbf{Z} . Each element can either contribute positively to the difference – meaning that using the GTAP element rather than the WIOD element makes the difference positive, or negatively – meaning that using the GTAP element rather than the WIOD element makes the difference negative. Both \mathbf{Z} and \mathbf{y} , in this case, contribute towards the negative difference, whereas the inverse output has a positive effect. The emissions vector \mathbf{f} makes little difference in this case. The overall difference of -14.3 MtCO₂ is the sum of the positive and negative differences and is therefore the net difference between the paths. The percentage values in each row calculate the influence each element has on the gross difference. The first row of Table 5 is the path representing UK transport emissions in transport products and here the difference is positive, meaning that GTAP's path is higher than WIOD and the majority of the difference (41%) is due to the final demand element in GTAP being larger than the element in WIOD.

3.3. Global results

The UK case study was used to explain how results were generated and to give an example of how to interpret the findings. This study has calculated the structural paths for the 40 common countries from Eora CC, EXIOBASE CC, GTAP CC and WIOD CC.

Table 6 summarises the characteristics of the 100 largest paths from each database. The sum of the largest 100 paths is largest in Eora and a greater proportion of the total is made up of very large (> 500 MtCO₂) paths in the Eora database. In each database, zero-th order paths make up the majority of the total and paths originating from the USA contribute to just under half of the sum of the largest 100 paths. There is some disagreement in the proportion of paths originating in different industrial sectors. Eora and GTAP show 26% and 22% of paths originating in the transport sector, where as EXIOBASE and WIOD estimate 8% and 7%. Eora reports a lower proportion of paths originating in the electricity, gas and water sector and a larger proportion in the construction sector.

Table 4. Elements in the top 20 PD.

Rank	GTAP						WIOD							
	f	x^{-1}	Z	x^{-1}	Z	y	f	x^{-1}	Z	x^{-1}	Z	y		
1	131.4	3.35E-6				82,228	91,871	4.27E-6				53,326		
2	201.0	1.09E-5	6,994	1.09E-5		31,883	180,881	5.89E-6	46,163	5.89E-6		67,862		
3	582.8	8.13E-7				5,075	974,971	1.33E-6				10,028		
4	6.4	9.93E-7				810,089	20,297	9.07E-7				840,407		
5	201.0	1.09E-5	11,139	9.93E-7		810,089	180,881	5.89E-6	11,716	9.07E-7		840,407		
6	1.8	1.94E-6				418,517	15,466	140E-6				447,826		
7	201.0	1.09E-6	6,994	1.09E-5	6,994	1.09E-5	31,883	180,881	5.89E-6	46,163	5.89E-6	46,163	5.89E-6	67,862
8	2.8	2.55E-6				244,393	10,438	2.58E-6				239,438		
9	9.1	8.26E-6				17,387	28,772	1.11E-5				17,807		
10	201.0	1.09E-5	3583	5.35E-6		128,663	180,881	5.89E-6	3,506	7.42E-6		50,191		
11	582,8	8.13E-7	1850	9.93E-7		810,089	974,971	1.33E-6	4,745	9.07E-7		840,407		
12	949.1	1.07E-6				3877	758,498	1.29E-6				300		
13	27.8	7.42E-6				2354	158,643	9.56E-6				2575		
14	131.4	3.35E-6	11,003	5.35E-6		128,663	91,871	4.27E-6	4,072	7.42E-6		50,191		
15	201.0	1.09E-5	6530	8.47E-7		366,039	180,881	5.89E-6	5,227	7.23E-7		495,003		
16	201.0	1.09E-5				31,883	180,881	5.89E-6				67,862		
17	131.4	3.35E-6	47,889	3.35E-6		82,228	91,971	4.57E-6	38,755	4.27E-6		53,326		
18	1,224	3.10E-5				5802	5,738	2.00E-5				21,718		
19	6,897	1.67E-5	2588	1.09E-5		31,883	22,185	1.14E-5	23,273	5.89E-6		67,862		
20	201.0	1.09E-5	9279	3.62E-6		10,850	180,881	5.89E-6	7,394	4.65E-6		28,741		

Table 5. SPDA results for UK GTAP and WIOD largest PD.

Rank	f effect	x^{-1} effect	Z effect	x^{-1} effect	Z effect	x^{-1} effect	y effect	Diff MtCO ²
1	10.1 34%	-7.1 24%					12.2 41%	15.3
2	1.7 3%	10.3 17%	-25.3 43%	10.3 17%			-11.3 19%	-14.3
3	-3.3 31%	-3.2 30%					-4.2 40%	-10.6
4	-10.9 89%	0.9 8%					-0.4 3%	-10.3
5	1.5 12%	8.6 68%	-0.7 6%	1.3 10%			-0.5 4%	10.1
6	-9.9 80%	2.0 16%					-0.4 3%	-8.3
7	0.4 2%	2.4 12%	-5.0 25%	2.4 12%	-5.0 25%	2.4 12%	-2.4 12%	-4.9
8	-4.7 97%	-0.0 1%					0.1 2%	-4.7
9	-3.4 77%	-0.9 22%					-0.1 2%	-4.4
10	0.3 6%	1.9 31%	0.1 1%	-1.1 18%			2.8 45%	4.0
11	-1.1 25%	-1.1 24%	-1.9 44%	0.2 5%			-0.1 2%	-4.0
12	0.5 10%	-0.4 9%					3.6 81%	3.6
13	-2.7 80%	-0.5 15%					-0.2 5%	-3.4
14	0.6 13%	-0.5 10%	1.6 33%	-0.6 13%			1.6 32%	2.7
15	0.3 8%	1.9 43%	0.7 16%	0.5 11%			-1.0 22%	2.4
16	8.1 7%	47.5 42%					-58.0 51%	-2.4
17	1.7 24%	-1.2 17%	1.0 14%	-1.2 17%			2.0 29%	2.3
18	-1.5 44%	0.6 17%					-1.3 39%	-2.3
19	-1.1 23%	0.5 10%	-1.7 35%	0.8 17%			-0.8 16%	-2.3
20	0.2 7%	1.1 42%	0.4 15%	-0.5 17%			0.7 23%	2.0

Next we calculate the PD for the six database pairings:

- Eora CC vs. EXIOBASE CC
- Eora CC vs. GTAP CC
- Eora CC vs. WIOD CC
- EXIOBASE CC vs. GTAP CC
- EXIOBASE CC vs. WIOD CC
- GTAP CC vs. WIOD CC

This paper summarises the results by means of a series of questions:

- Is one database consistently calculating the larger of the paths when two databases are compared?
- What is the frequency distribution by size of path difference?

Table 6. Characteristics of the 100 largest paths in each MRIO database.

Characteristic	Eora		EXIOBASE		GTAP	WIOD		
Sum of largest 100 paths (MtCO ₂)	8803	6462	6313	6107		8803	6462	6313
					Sum of largest 100 paths (MtCO ₂)			
<i>Sum of paths by size (MtCO₂)</i>								
500 < x	3596	(41%)	1104	(17%)	812	(13%)	1115	(18%)
100 < x ≤ 500	1706	(19%)	2256	(35%)	2076	(33%)	1656	(27%)
50 < x ≤ 100	1517	(17%)	903	(14%)	1332	(21%)	1168	(19%)
20 < x ≤ 50	1808	(21%)	2048	(32%)	1799	(28%)	1961	(32%)
x ≤ 20	176	(2%)	150	(2%)	294	(5%)	206	(3%)
<i>Sum of paths by path order (MtCO₂)</i>								
0	6676	(76%)	4110	(64%)	3856	(61%)	4037	(66%)
1	1932	(22%)	1955	(30%)	1990	(32%)	1722	(28%)
2	155	(2%)	353	(5%)	406	(6%)	282	(5%)
3	40	(0%)	43	(1%)	61	(1%)	67	(1%)
<i>Sum of paths by source region (MtCO₂)</i>								
USA	4119	(47%)	2984	(46%)	2569	(41%)	2753	(45%)
China	1793	(20%)	1570	(24%)	1583	(25%)	1367	(22%)
Russia	701	(8%)	289	(4%)	382	(6%)	346	(6%)
India	528	(6%)	294	(5%)	396	(6%)	332	(5%)
Other	1663	(19%)	1324	(20%)	1382	(22%)	1308	(21%)
<i>Sum of paths by source industry (MtCO₂)</i>								
Electricity, gas and water	4084	(46%)	4126	(64%)	4073	(65%)	3623	(59%)
Transport	2280	(26%)	508	(8%)	1393	(22%)	432	(7%)
Construction	1378	(16%)	147	(2%)	57	(1%)	71	(1%)
Petroleum, chemicals and other non-metallic minerals	259	(3%)	485	(8%)	318	(5%)	642	(11%)
Other	801	(9%)	1196	(19%)	476	(8%)	1337	(22%)

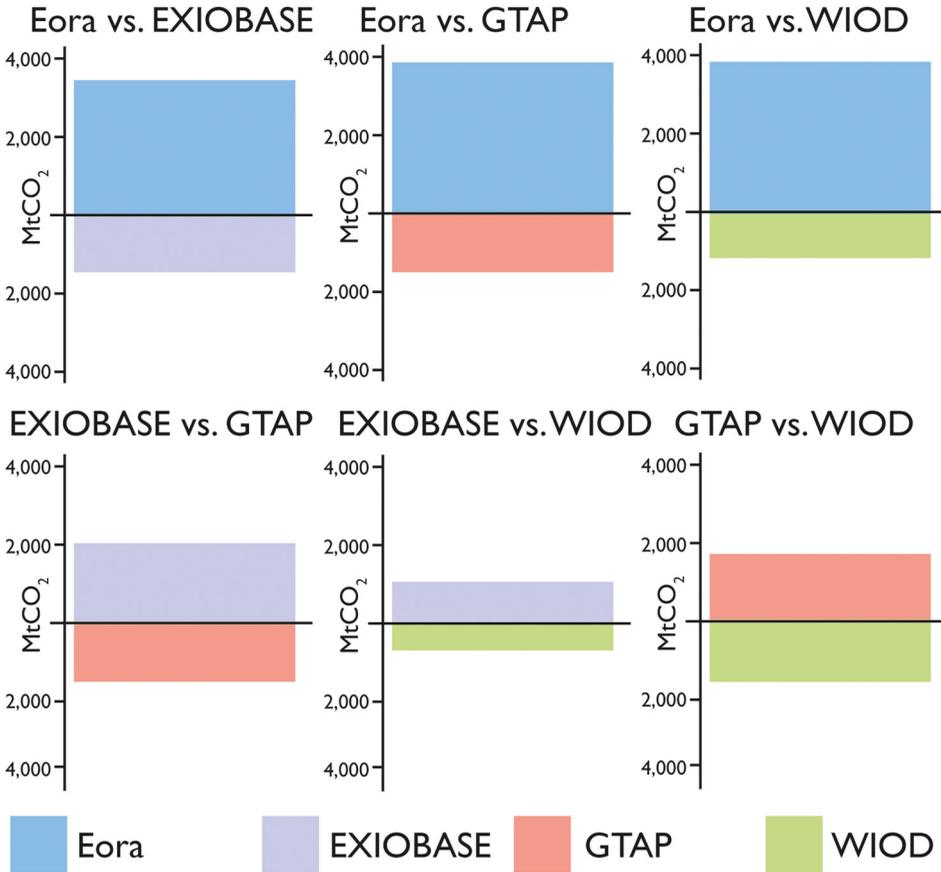
- What orders of paths contribute to the total of the top 100 PD?
- Are there particular countries that tend to produce large PD?
- Are there particular sectors that tend to produce large PD?
- Are there particular elements within the series expansion equation that tend to be responsible for most of the difference between paths?
- In what type of paths does the emissions data contribute most to the difference?
- In what type of paths does the monetary data contribute most to the difference?

3.3.1. Is one database consistently calculating the larger of the paths when two databases are compared?

In general, Eora estimates CBAs to be larger than the estimates from EXIOBASE, GTAP and WIOD (see Figure 1). This finding is also demonstrated in the SPA where Eora paths tend to be larger than their counterparts in EXIOBASE, GTAP and WIOD. Figure 2 shows that from the sum of the top 100 PD, Eora paths contribute 72% of the gross difference when paired with EXIOBASE, 74% when paired with GTAP and 75% when paired with WIOD. EXIOBASE paths make up slightly more of the difference when paired with GTAP and WIOD. GTAP and WIOD paths share the difference almost equally.

In the following questions, the difference assigned to each pairing is disaggregated to show characteristics of the difference. Figures 3–6 display the findings and Table 3 in the SI gives the full results

Figure 2. (Colour online) Sum of top 100 PD disaggregated by the database which exhibits the larger of the paired paths.



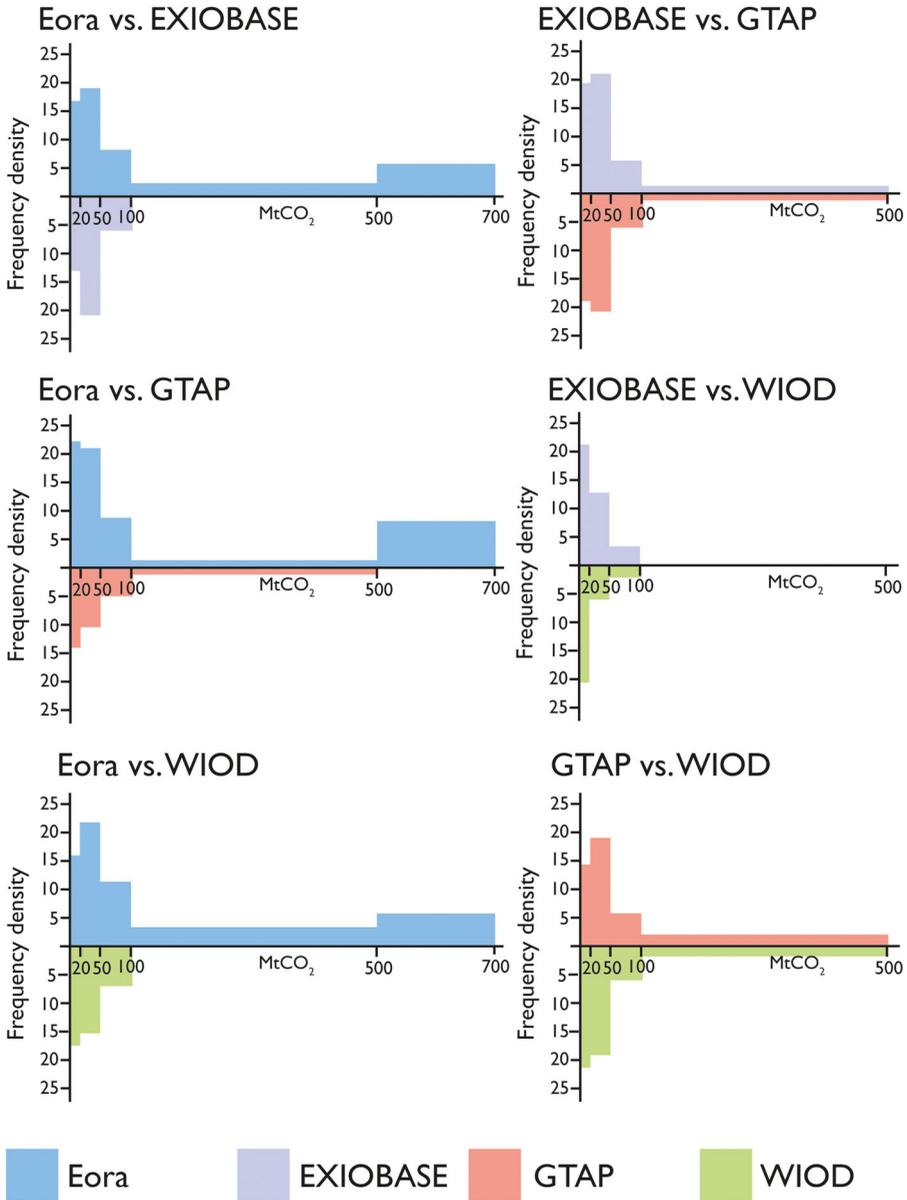
3.3.2. What is the frequency distribution by size of path difference?

Each database pairing contains a small number of very large PD and the majority of the PD are between 10 and 20 MtCO₂. When Eora and EXIOBASE are compared, (see Figure 3, top left), we find two paths that differ by more than 500 MtCO₂. This contributes 1,225 MtCO₂ (26%) to the gross difference. To put this into context, the United Nations (UNFCCC, 2007) reports global CO₂ emissions to be 30,113 MtCO₂. 1225 MtCO₂ represents 4% of the global total. Paths where Eora reports highest are the only ones where differences of over 500 MtCO₂ are observed. This finding is consistent with the character of the Eora paths shown in Table 6. This also reinforces the conclusions drawn in Owen et al., (2014) that GTAP and WIOD are more similar to each other than when either database is paired with Eora. EXIOBASE and WIOD have the fewest very large PD with just over 50% of the gross difference from PD of over 20 MtCO₂.

3.3.3. What orders of paths contribute to the total of the top 100 PD?

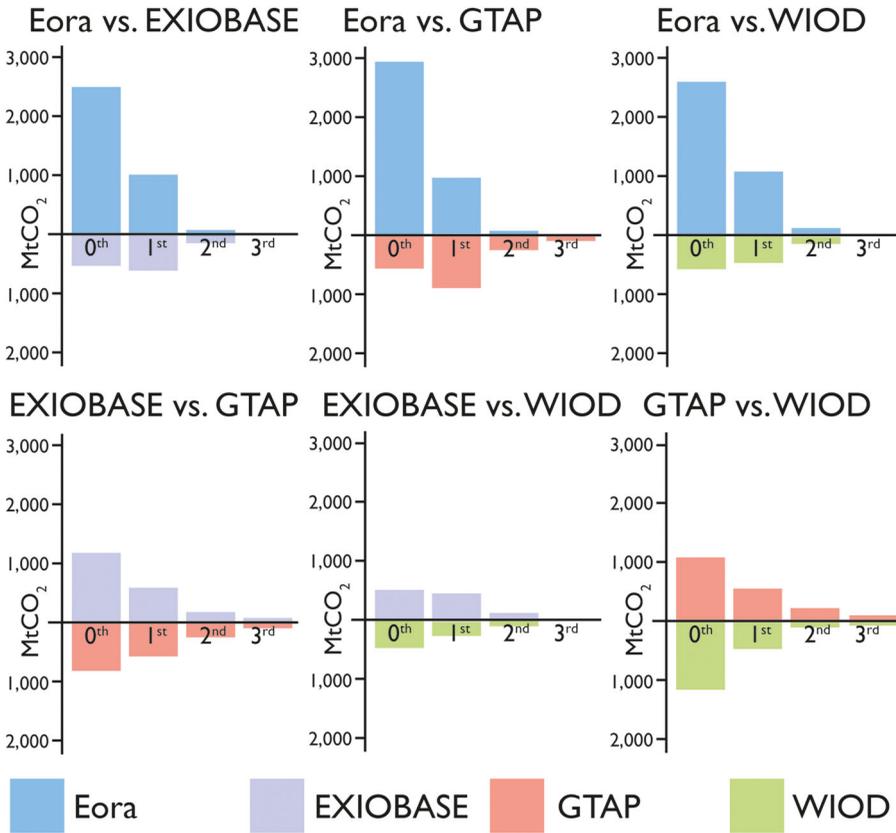
In all six pairings, the majority of gross difference is from zero-th order paths as shown in Figure 4. These are paths from the source emissions straight to final demand of the

Figure 3. (Colour online) Sum of top 100 PD by size of path difference and disaggregated by the database which exhibits the larger of the paired paths.



same product, by-passing the interactions matrix Z . This means that the cause of the difference must lie in the emissions vector f , the output vector x and final demand vector y . In the Eora and EXIOBASE SPD comparison, 97% of the difference is in zero-th and first order paths. For Eora and GTAP this figure is 96%, Eora and WIOD 97%, EXIOBASE and GTAP 91%, EXIOBASE and WIOD 95% and finally for GTAP and WIOD, 93%. Only pairings involving GTAP have PD that are third order in the top 100. To contain a third

Figure 4. (Colour online) Sum of top 100 PD by path order and disaggregated by the database which exhibits the larger of the paired Paths.

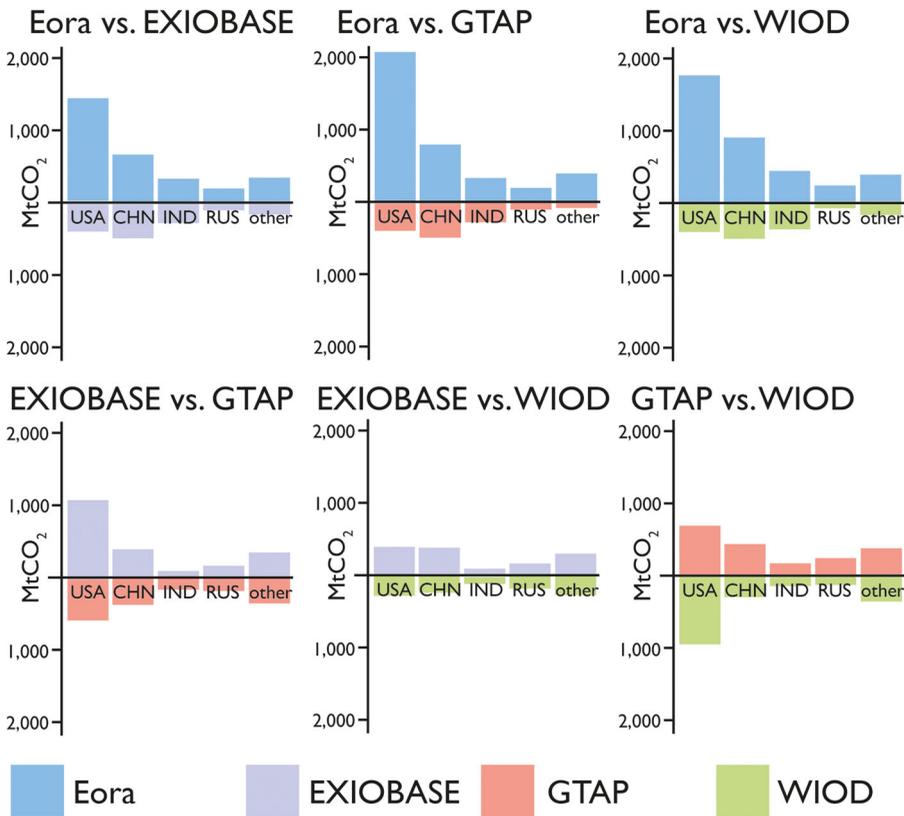


order path in the top 100 differences means that there is likely to be large differences in the Z matrix.

3.3.4. Are there particular countries that tend to produce large PD?

There are no paths in the top 100 PD for any of the six pairings where the path crosses a country border. All paths with large PD are contained within a single country. This is not surprising since none of the 100 largest paths identified by the SPA cross borders. Figure 5 shows that for every pairing, the gross difference is made up of paths from the USA, followed by China, India and Russia. These four nations make up 87%, 90%, 88%, 80%, 76% and 80% of the gross difference from the top 100 PD from the Eora and EXIOBASE; Eora and GTAP; Eora and WIOD; EXIOBASE and GTAP; EXIOBASE and WIOD; and GTAP and WIOD SPA calculations, respectively. Interestingly, Table 6 shows that paths originating in India contribute to 5–6% of the sum of the 100 largest paths in each of the databases, but paths originating in India contribute to 14% of the gross difference when Eora and EXIOBASE are paired, 11% for Eora and GTAP and 15% for Eora and WIOD. Figure 5 also reveals that it is not the case that one database consistently calculates larger paths for those originating in India indicating that the difference is unlikely to be caused by one database simply containing larger values.

Figure 5. (Colour online) Sum of top 100 PD by country of origin and disaggregated by the database which exhibits the larger of the paired Paths.



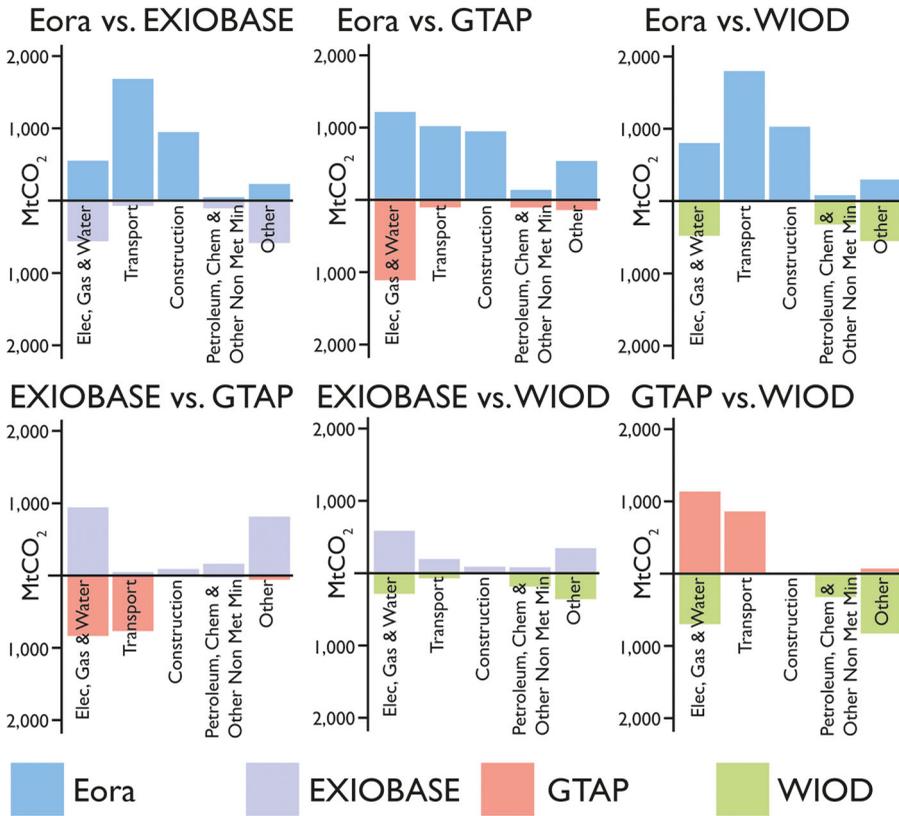
3.3.5. Are there particular sectors that tend to produce large PD?

When comparing Eora and EXIOBASE, Figure 6 shows that 36% of the gross difference is from paths which originate in the transport sector, with Eora reporting higher paths in most cases. The electricity, gas and water sector; petroleum, chemicals and non-metallic mineral sector; and construction sector also feature heavily in the contribution to the overall difference. PD involving GTAP feature the electricity, gas and water sector for a larger proportion of the paths than the Eora and EXIOBASE or Eora and WIOD pairings. Interestingly, Table 6 shows that transport is the origin sector of 7% of the 100 largest paths in WIOD and 26% in Eora, yet for the paths with the largest difference, we find transport originating paths contribute to 35% of the gross difference between Eora and WIOD, with Eora always reporting the larger path. Transport does not feature in many of the paths with high differences when comparing EXIOBASE and WIOD. This finding is discussed later.

3.3.6. Are there particular elements within the series expansion that tend to be responsible for most of the difference between paths?

SPD allows us to identify the contribution towards the difference that each element in the path makes. To summarise the information we first consider which element contributes most to the path difference. Figure 7 shows, that for the gross difference from the top 100

Figure 6. (Colour online) Sum of top 100 PD by sector of origin and disaggregated by the database which exhibits the larger of the paired Paths.



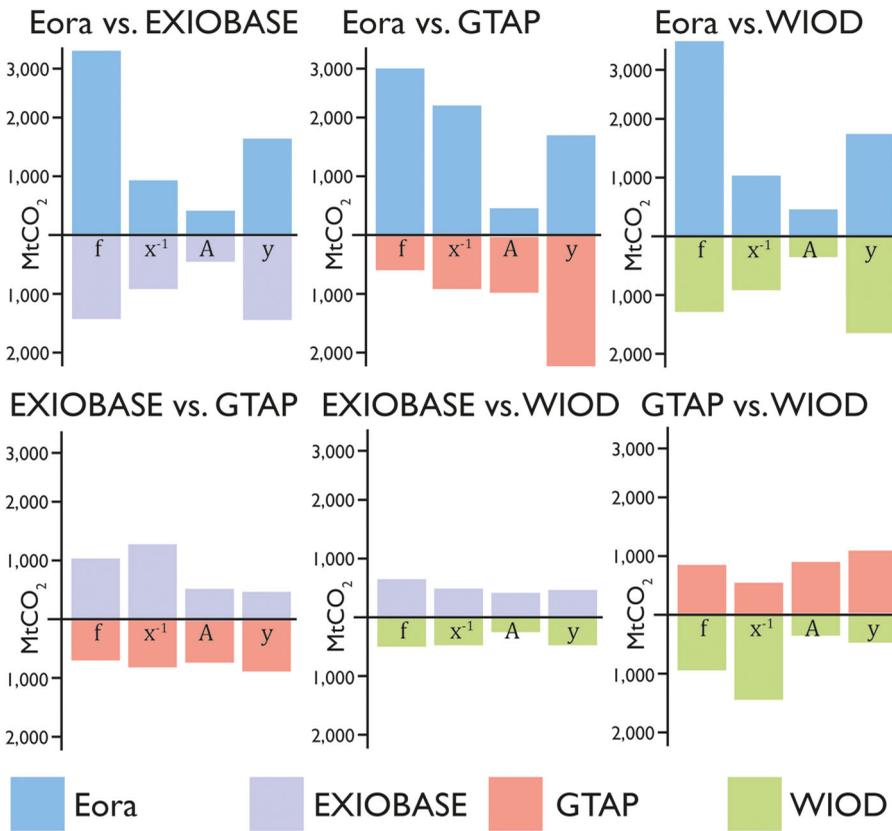
PD between the Eora and EXIOBASE databases, the element from the emissions vector is the largest contributor of difference (45%). The final demand figure contributes 30% of the differences, followed by total output (17%) and the element in the transaction matrix (8%). The emissions vector (f) contributes most to the difference in pairings involving Eora and the transactions matrix (A) contributes most to the difference in pairings involving GTAP.

3.3.7. What are the characteristics of paths where the emissions or the monetary data contribute most to the difference?

Finally, we characterise the types of paths where emissions are the causes of difference and the types of paths where the monetary information is the cause of difference. Table 7 shows the top 10 paths where the element in the emissions vector was the largest contributor to the difference. The effect of each element is shown in the adjacent ‘ $\Delta MtCO_2$ ’ column. We find that the transport (TRNS), construction (CNST), trade (TRAD) and public administration, education, health and defence sectors (PAEH) are where the emissions vectors disagree. Surprisingly, the electricity, water and gas (ELGW) sector does not appear high in the list of paths where the emissions contribution differs substantially.

The three Eora pairings show very large PD where the value chain starts with the emissions from USA TRNS and China CNST. Paths starting with China CNST do not appear

Figure 7. (Colour online) Sum of top 100 PD by equation element responsible for the largest portion of the differences the disaggregated by the database which exhibits the larger of the paired paths.



in the remaining pairings which suggests that Eora may overestimate the emissions from this sector.

Table 8 shows the top 10 paths where either total output, the transaction matrix or the final demand matrix were the highest contributors towards the path difference. Emissions for the ELGW sector seem to align between databases, but the monetary data differs quite staggeringly and is one of the major contributors towards PD. The three GTAP pairings show very large PD where the value chain involves a transaction with the ELGW sector. Compared to Eora, EXIOBASE and WIOD, GTAP underestimates the monetary information for the USA ELGW sector and over estimates for China's ELGW sector.

4. Discussion

4.1. Domestic value chains

In the top one hundred paths with the largest differences, every path from every database pairing is entirely contained within one single country. There are no paths with very large differences that describe imports to final or intermediate demand. Emissions in trade account for around one quarter of global emissions (Davis and Caldeira, 2010; Peters et al., 2012). but the 'off-diagonal' blocks within MRIO databases which show the imports to

Table 7. Top 10 PD where the emissions element is the largest contributor to the overall difference.

	Eora and EXIOBASE	Δ Mt CO ₂	Eora and GTAP	Δ Mt CO ₂	Eora and WIOD	Δ Mt CO ₂	EXIOBASE and GTAP	Δ Mt CO ₂	EXIOBASE and WIOD	Δ Mt CO ₂	GTAP and WIOD	Δ Mt CO ₂
1	USA TRNS	631	CHN CNST	604	USA TRNS	659	USA PAEH	175	USA PAEH	-81	USA PAEH	-258
2	CHN CNST	594	USA TRNS	564	CHN CNST	597	CHN METP > CHN CNST	77	USA BSNS	66	USA TRAD	-108
3	USA TRNS > USA PAEH	301	USA PAEH	134	USA TRNS > USA PAEH	295	USA TRAD	71	JPN ELGW	62	USA TRNS	95
4	IN CNST > IND TRNS	116	USA TRNS > USA PAEH	120	USA PAEH	-123	USA TRNS	-67	CHN METP > CHN CNST	46	USA BSNS	-56
5	CHN METP > CHN CNST	-112	INS CNST > IND TRNS	115	IND CNST > IND TRNS	115	USA CNST	49	CHN PETC > CHN CNST	-45	CHN PETC > CHN CNST	-49
6	MEX TRNS	88	USA TRNS > USA TRAD	80	USA TRAD	-100	JPN ELGW	42	USA POST	24	USA PETC	-41
7	USA TRNS > USA TRAD	86	USA BSNS	62	CHN PETC > CHN CNST	-98	MEX TRNS	-41	CHN METP > CHN ELGW	24	USA CNST	-36
8	USA TRAD	-68	IND ELGW > IND AGRI	-57	MEX TRNS	82	FRA TRNS	-38	JPN PAEH	-21	DEU TRNS	36
9	IND CNST > IND BSNS	68	USA TREQ	55	USA TRNS > USA TRAD	73	USA POST	35	DEU ELGW	-21	MEX TRNS	36
10	USA BSNS	-60	USA CNST	54	IND CNST > INS BSNS	68	USA BSNS > USA PAEH	32	RUS PETC RUS CNST	-19	FRA TRNS	35

Table 8. Top 10 PD where elements from the total output vector, the transaction matrix or the final demand vector is the largest contributor to the overall difference.

	Eora and EXIOBASE	Δ Mt CO ₂	Eora and GTAP	Δ Mt CO ₂	Eora and WIOD	Δ Mt CO ₂	EXIOBASE and GTAP	Δ Mt CO ₂	EXIOBASE and WIOD	Δ MtCO ₂	GTAP and WIOD	Δ Mt CO ₂
1	USA ELGW	393	USA ELGW	685	USA ELGW	382	USA ELGW	292	CHN ELGW	126	USA ELGW	-303
2	IND CNST	116	CHN ELGW	-180	IND CNST	112	USA TRNS > USA PAEH	-181	RUS ELGW > RUS PAEH	47	CHN ELGW	285
3	USA ELGW > USA TRAD	-75	RUS ELGE	-159	CHN ELGW	103	CHN ELGW	-158	USA ELGW > USA PAEH	47	USA TRNS > USA PAEH	176
4	IND ELGW	-58	IND CNST	119	IND ELGW	-880	RUS ELGW	-137	CHN ELGW > CHN ELGW	47	USA ELGW > USA PAEH	154
5	CHNELGW > CHN ELGW	-46	USA ELGW > USA PAEH	-166	RUS ELGW > RUS PAEH	75	USA BSNS	121	USA ELGW > USA TRAD	45	RUS ELGW	153
6	IND ELGW > IND CNST	-44	IND ELGW	-111	CHN ELGW > CHN PAEH	-40	USA ELGW > USA PAEH	-107	USA TRAD	-31	USA ELGW > USA ELGW	89
7	CHN ELGW > CHN PAEH	-36	USA ELGW > USA ELGW	-89	IND ELGW > IND CNST	-40	USA ELGW > USA ELGW	-87	USA ELGW > USA BSNS	29	CHN ELGW > CHN CNST	-65
8	RoW PETC	-29	US ELGW > RUS PAEH	87	USA ELGW > USA PAEH	37	RUS ELGW > RUS PAEH	60	USA TRNS	28	USA ELGW > USA TRAD	56
9	RUS ELGW > RUS ELGW	27	USA ELGW > USA TRAD	-86	USA ELGW > USA TRAD	-30	USA ELGW > USA BSNS	60	USA MINQ	-24	DEU ELGW	-45
10	USA PETC > USA CNST	-26	USA PETC	64	RoW PETC	-29	CHN ELGW > CHN ELGW	54	KOR ELGW	23	CHN TRNS	43

intermediate and final demand are often estimated based on proportionality assumptions (Bouwmeester and Oosterhaven, 2007; Tukker et al., 2009; Erumban et al., 2011; Peters et al., 2011). Owen et al. (2015) demonstrate that the sections of the transactions matrix Z that represent imports align less between databases than the domestic transactions but the authors also find that the effect of the difference in the import blocks on the overall difference in CBA is not as significant as that of other factors such as total emissions.

Some individual country level results do show paths that contain imports as having large differences for nations that rely on traded goods and Table 5, which shows the largest PD for the UK using GTAP and WIOD, has several such paths. However, the nations that have the largest emissions CBAs and the largest individual emissions supply chains tend to be countries like the USA, China and Russia that are not overly reliant on traded goods for intermediate and final demand. In addition, the largest paths often involve electricity, water and gas which are more likely to be domestically sourced.

4.2. Sources of difference from the emissions vector

Owen et al. (2014) concludes that differences in the emissions vector are a major cause of difference between Eora and GTAP and Eora and WIOD. Similarly, Moran and Wood, (2014) find that harmonising the emissions vector causes CBA calculated using Eora, EXIOBASE, GTAP and WIOD to converge. This study finds that the emissions element is the greatest cause of difference in 63 out of the top 100 paths with large differences between Eora and WIOD, contributing to 43% of the sum of the gross difference and 61 of the top 100 paths when Eora and EXIOBASE are compared (45% of the gross difference sum). Table 9 reveals that the total global emissions differ between databases with Eora estimating total Global emissions to be slightly larger than EXIOBASE, GTAP and WIOD. In particular, Eora's industrial emissions figure is considerably larger than that used in the other databases and the figure assigned to household direct emissions from home heating and private transportation is lower.

In this study, the SPD calculations focus on the difference in supply chains between the four MRIO databases and hence exclusively use the industrial emissions figures. And, because Eora's industrial emissions figure is so much larger than the amount used by the other three databases, it is unsurprising that Eora's emissions are a large source of difference.

There are two main reasons for databases attributing different proportions of total emissions to industries rather than households: (1) due to whether the residence or territorial principle is used for emissions allocation; and (2) the allocation of the road transport activity to users. The residence principle is used in a national accounting framework and states that emissions activity of a resident unit (i.e. a person or company) are allocated to

Table 9. Comparison of total global CO₂ emissions in Eora, EXIOBASE, GTAP and WIOD for the year 2007.

	Eora	EXIOBASE	GTAP	WIOD
Total global emissions 2007 (MtCO ₂)	30,431	28,975	26,524	29,218
Industrial	28,237	24,757	22,800	25,261
Household	2194	4218	3724	3957

the territory of residence (Genty et al., 2012). This means specifically that when calculating a national account, activities of tourists are removed and reallocated to the country of residence of the tourist and any domestic residents' activities abroad are added. This affects the overall level of the transport component in each country. The territorial principle allocates emissions to the country where they take place and are used in national statistics. The second affect relates to the distribution of road transport (classified as an own activity in energy balances) is allocated to industries and households undertaking the activity (as in energy accounts). In energy balances published by each country or international institutes such as the International Energy Agency (IEA), the road transport is represented as a single activity. Such a representation is not consistent with the industry/household delineation in National Accounts, where the emission should be recorded under the establishment/household undertaking the activity (Eurostat, 2014; UNDESA, 2015).

Both WIOD and EXIOBASE use the residence principle, and allocate road transport to the user (Genty et al., 2012; Tukker et al., 2013) and Figure 6 shows that the transport industry is not as significant a source of difference between EXIOBASE and WIOD as it is in any of the other five database pairings. In addition, the EXIOBASE and WIOD column in Table 7, which shows the top PD where the emissions vector is the most significant cause of difference, does not have a single path containing the transport sector. Eora does not correct for the residence principle or road transport (Lenzen, Kanemoto, et al., 2012), which may explain why the allocation to households and industry is different in Table 9 and why transport appears prominently in both Table 7 and Figure 6 where Eora is compared to EXIOBASE and WIOD. GTAP also uses a territorial treatment but allocates international transportation to consumers not producers (Peters et al., 2011). GTAP v6 used a 50/50 (gasoline) and a 75/25 (diesel) split for fuels used in road transport to the road transport industry and to households respectively, ignoring usage by other industries (McDougall and Lee, 2006).

4.3. Sources of difference from the monetary data

This study finds that the majority of the difference in paths, where the monetary data is the largest contributor towards the overall path difference, involves the electricity, gas and water sector. Either the total output, the element from the transactions matrix (Z) or the final demand figure for this sector is very different between the databases. In Table 10 we show the proportion of the electricity, gas and water production mix for each country in the CC that is supplied by that sector itself (including intermediate imports of electricity, gas and water from abroad), taking the values from the Z matrix for each database.

Table 10 shows that there is a large difference in the electricity, gas and water proportion across the databases and this discrepancy was brought to our attention by the SPA. There are a number of reasons as to why the monetary data could differ for this sector. The definition of what is included as electricity, gas and water may be different for the different databases. For example, in some countries gas manufacturing is treated as a margin sector.

In general, Eora, EXIOBASE and WIOD agree on the electricity, gas and water proportions and GTAP is the outlier. However when one reads down Table 10, the proportions vary significantly between countries. Eora, EXIOBASE and WIOD report values of over 40% for Austria, Brazil, the UK, Portugal, Slovakia and Turkey but less than 1% for Canada.

Table 10. Proportion of the electricity, gas and water sector that electricity, gas and water supplies.

		Eora_CC (%)	EXIOBASE_CC (%)	GTAP_CC (%)	WIOD_CC (%)
1	AUS	11	25	14	18
2	AUT	63	73	31	73
3	BEL	26	24	11	24
4	BLG	20	12	27	12
5	BRA	57	0	38	46
6	CAN	0	1	11	0
7	CHN	24	47	15	43
8	CYP	7	29	9	5
9	CZE	56	30	17	30
10	DEU	22	30	14	30
11	DNK	14	2	11	7
12	ESP	27	34	9	35
13	EST	15	20	14	20
14	FIN	5	8	6	8
15	FRA	24	37	13	37
16	GBR	50	41	19	44
17	GRC	13	30	18	29
18	HUN	17	14	19	15
19	IDN	11	22	19	18
20	IND	22	37	19	32
21	IRL	42	54	8	60
22	ITA	27	24	11	24
23	JPN	13	11	11	10
24	KOR	24	25	13	29
25	LTU	16	20	18	20
26	LUX	2	39	52	29
27	LVA	21	49	14	19
28	MEX	25	25	32	26
29	MLT	7	64	10	15
30	NLD	39	36	33	37
31	POL	12	6	13	7
32	PRT	61	61	9	66
33	ROU	44	29	31	29
34	RUS	25	21	16	9
35	SVK	61	46	21	46
36	SVN	14	32	14	33
37	SWE	13	15	9	16
38	TUR	51	58	12	44
39	TWN	14	17	11	15
40	USA	0	1	21	0
41	ROW	26	28	23	19

These differences can be interpreted a number of ways – in countries such as the USA, a single entity is likely modelled as the generator and distributor (no sale between entities), whilst in countries such as Austria, a sale between generator and distributor likely occurs. Further complicating the issue is countries such as Australia, which are adopting progressive classification of distribution and transmission as margin commodities in order to separate the basic good from the delivery. Clearly it would be preferable if a consistent definition was used across all countries in MRIO databases. As WIOD and EXIOBASE seek to fully respect National Statistical office data, this raises the question in the short term of whether this is the best approach, and in the long term if umbrella statistical agencies such as Eurostat or efforts by the OECD and UN are able to further harmonise these efforts. Finally, it can be further be noted, that whilst large intra-sectorial flows show strong differences at the path level for the reasons above, the final emissions associated with products

going to final demand is insensitive to the value of the intra-sectorial flow (e.g. see Miller and Blair 2009, p. 279).

Apart from the intra-sectorial flows, electricity is a sector further fraught with difficulty when monetary data is used to describe the distribution of electricity use. Different industrial sectors spend different amounts of money to receive the same KWh of electricity because the price per KWh differs by sector. GTAP does not rely on user submitted values in the energy rows of the I–O tables. Here physical data on energy use in Joules is taken from the IEA, converted to monetary values and placed in the I–O tables (Peters et al., 2011). This removes the problem of electricity prices described.

In Table 11, we calculate, from the **Z** matrix, the proportion of domestic electricity, gas and water supply that goes to each other domestic sector, including domestic final demand. Electricity, gas and water own use is zeroed, to highlight the proportional allocation of impact to consumers, and exports to intermediate and final demand in other countries are combined in the export column. Table 11 allows us to compare the proportional use by selected countries for all four MRIO databases. If the proportional use deviates from the multi-model mean by more than one standard deviation, the value is bold and shaded. Table 11 shows that for large developed nations, such as Australia, Canada, Germany, Spain, the UK and the US, GTAP allocates a greater share of electricity, gas and water to manufacturing (MANU) and a smaller share to domestic final demand (DOM FD). This implies that the price per unit for manufacturing sectors is lower than average and Eora, EXIOBASE and WIOD thus underestimate the energy use in these sectors. It is not conclusive whether this pattern holds for China, India and Russia because there are larger overall differences between the four MRIO databases for these nations. Further investigation into this re-proportioning construction method is needed before we can conclusively state that the electricity, gas and water sector is more reliably described in GTAP.

4.4. Future construction of MRIO databases

4.4.1. Residence or territorial principle?

This study recommends using the residence principle for allocation within the emissions vector. The residence principle is the technique used within the system of national accounts, thus this should be reflected in the data used to construct consumption accounts. It is perhaps an indication of the intention of the MRIO database construction community that the most recent database, EXIOBASE takes the residence principle (Usubiaga and Acosta-Fernández, 2015).

4.4.2. Economic data as a proxy for physical flows

An interesting finding is that GTAP's method of reallocating electricity spends to match the energy used proportions has significant difference when comparing structural paths. The investigations in this study have highlighted several classic issues in I–O analysis. If the energy sector covers both the energy producing and distributing functions of energy supply then this is an example of allocation uncertainty issue identified by Lenzen (2000).

Lenzen (2000) warns of proportionality assumption uncertainties explaining that when monetary data is used in I–O tables to represent a physical flow of commodities between industries one assumes that a dollar spend on energy by the energy sector is the same

Table 11. Proportion of electricity, water and gas supplied to each sector (excluding electricity, water and gas own use) by model and for selected countries.

		AGRI and MINQ (%)	MANU (%)	CNST (%)	TRNS (%)	TRAD (%)	SERV (%)	DOM FD (%)	EXPORT (%)
AUS	Eora	6	18	1	4	13	18	38	1
	E'BASE	8	17	5	3	8	19	37	2
	GTAP	9	32	1	3	7	17	31	1
	WIOD	6	13	5	4	7	19	45	0
CAN	Eora	5	21	0	1	10	19	40	5
	E'BASE	5	19	0	2	7	20	37	9
	GTAP	10	28	0	2	7	17	28	7
	WIOD	6	20	1	1	7	19	38	8
CHN	Eora	10	66	4	2	0	9	10	0
	E'BASE	12	57	4	2	4	8	13	0
	GTAP	8	64	1	1	3	7	16	1
	WIOD	12	61	5	2	4	9	7	0
DEU	Eora	2	37	0	2	7	13	34	5
	E'BASE	2	24	2	2	5	13	45	9
	GTAP	3	39	0	3	2	19	29	5
	WIOD	2	25	1	2	5	12	42	11
ESP	Eora	4	30	2	3	7	21	27	1
	E'BASE	3	28	2	4	5	21	27	1
	GTAP	4	33	1	2	2	22	26	3
	WIOD	3	28	2	3	5	19	29	1
FRA	Eora	2	21	1	1	6	18	45	5
	E'BASE	2	21	1	2	6	17	47	3
	GTAP	2	23	1	3	6	25	31	8
	WIOD	1	21	1	2	6	17	47	4
GBR	Eora	2	27	0	1	4	14	50	2
	E'BASE	3	20	1	2	5	14	55	0
	GTAP	2	31	1	3	2	22	38	1
	WIOD	2	19	1	1	5	15	55	1
IND	Eora	3	15	0	10	0	3	67	0
	E'BASE	3	28	2	4	14	21	27	1
	GTAP	4	33	1	2	7	22	26	3
	WIOD	3	28	2	3	14	19	29	1
RUS	Eora	8	39	3	7	4	27	11	1
	E'BASE	2	21	1	2	6	17	47	3
	GTAP	2	23	1	3	6	25	31	8
	WIOD	2	21	1	2	6	17	47	4
USA	Eora	2	11	0	1	4	18	64	1
	E'BASE	3	20	1	2	5	14	55	0
	GTAP	2	31	1	3	2	22	38	1
	WIOD	2	19	1	1	5	15	55	1

amount of energy as a dollar spend by the service sector. In reality, different industries pay different prices for energy and the treatment of margins is central here – greater harmonisation on how transmission and distribution costs are modelled so that basic price values of electricity generation only represent the physical flow of electricity would be preferable. Lenzen's suggested pragmatic solution is to replace entries with physical units. Dietzenbacher and Stage (2006) point out, however, that this hybrid solution, where an I–O table contains a mix of units, produces a database unsuitable for structural decomposition analyses. GTAP's solution of replacing spends with the monetary proportion of the actual energy used can be shown to inadvertently avoid the allocation uncertainty issue described above and goes some way to avoiding the proportionality assumption.

4.5. Future use of MRIO outcomes in policy

Whilst calculations of national CBA have been shown to be robust (Lenzen et al., 2010), this study indicates that calculations which involve extracting smaller portions of national level results table such as finding product footprint may not be as accurate. Wiedmann et al. (2011) explain that product footprints may become policy relevant if eco-labelling becomes a requirement of product sustainability standards. MRIO databases are less similar at this level of detail and the data is subject to higher levels of uncertainty due to the assumptions made in the database construction starting to have an effect at this scale. In this study, the most detailed level of data is explored; the value chain. Results show that there are large variations in the size of supply chains between databases. These differences obviously reflect the different source data used but choice of source data does not impede the recommendation for using MRIO database to assess global value chains – alternative source data can always be supplemented in the database. The effect of different construction techniques is more of a concern here. There is no set of agreed steps for constructing the emissions vectors; dealing with missing data; balancing the database; and thus each MRIO database has its own unique construction method. The findings from this study suggest that the choice of territorial or residence principle for generating the emissions vector and the technique used in GTAP for dealing with electricity price variations have large effects on the outcomes. It is therefore suggested that global value chain data is not yet robust enough to be used in climate policy. Nevertheless Lenzen, Kanemoto, et al. (2012) and Lenzen, Moran, et al. (2012) exploration of this approach shows its potential in demonstrating the interconnectedness of consumers, producers and associated environmental impacts in an increasingly globalised world.

4.6. Using aggregated data

The conclusions drawn are based on aggregated versions of the original MRIO databases. Owen et al. (2015) demonstrate that the aggregated versions are reasonable representations of the original databases using a series of matrix difference statistics. One could argue that care needs to be taken in interpreting results that are highly aggregated. However, this study calculated SPD on paths of length 11, which represented fourth order paths. Finding and identifying the fourth order paths from the original versions of the database would be a processing heavy calculation due to the sizes of the original matrices. The aggregated databases are quicker to use. Results using the aggregated versions could be seen as an initial sifting process. Now that the paths with cause for concern have been identified, the sectors involved could be studied in more detail at the disaggregated level.

5. Conclusion

This study represents the first time that SPD has been used with a S–S decomposition and is the first to compare PD between multiregional input–output (MRIO) databases. We show that SPD is an important technique for highlighting and explaining differences in the global value chains produced by MRIO databases in the calculation of consumption-based accounts. The work expands upon the findings from Owen et al. (2014) by including the EXIOBASE MRIO database and allowing consideration of difference at the sector level.

The findings presented in this paper will be of great interest to constructors of MRIO databases since they help to explain why MRIO database outcomes differ. The findings also point to key areas where harmonisation of source data and construction techniques could bring about convergence of results. For example, we find that sources of difference can be traced to the different emissions source data used, as well as choice of construction techniques, such as using the residence principle rather than the territorial for emissions allocation and redistributing energy spends based on physical data. This work should also be of interest to the users of MRIO database outcomes since it highlights at which scales results are most consistent. For example, national level consumption-based accounts are more robust than global value chain information.

We recommend that this work be extended to include future MRIO systems and to consider data from different years.

Acknowledgements

We would like to thank two anonymous reviewers whose insightful comments helped to improve the content of this paper.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

AO's time was supported by the research programme of the UK Energy Research Centre, supported by the UK Research Councils under [EPSRC award EP/L024756/1] and the RCUK Energy Program's funding for the Centre for Industrial Energy, Materials and Products [grant reference EP/N022645/1].

References

- Acquaye, A., T. Wiedmann, K. Feng, R.H. Crawford, J. Barrett, J. Kuylenstierna, . . . S. McQueen-Mason (2011) Identification of "Carbon Hot-Spots" and Quantification of GHG Intensities in the Biodiesel Supply Chain Using Hybrid LCA and Structural Path Analysis. *Environmental Science & Technology*, 45, 2471–2478.
- de Boer, P. (2009) Generalized Fisher Index or Siegel–Shapley Decomposition? *Energy Economics*, 31, 810–814.
- Bouwmeester, M. and J. Oosterhaven (2007) Technical Report: Inventory of Trade Data and Options for Creating Linkages. Report of the EXIOPOL Project.
- Crama, Y., J. Defourny and J. Gazon (1984) Structural Decomposition of Multipliers in Input–Output or Social Accounting Matrix Analysis. *Economie Appliquee*, 37, 215–222.
- Davis, S.J. and K. Caldeira (2010) Consumption-Based Accounting of CO₂ Emissions. *Proceedings of the National Academy of Sciences of the United States of America*, 107, 5687–5692.
- Defourny, J. and E. Thorbeck (1984) Structural Path Analysis and Multiplier Decomposition Within a Social Accounting Matrix. *The Economic Journal*, 94, 111–136.
- Dietzenbacher, E. and B. Los (1998) Structural Decomposition Techniques: Sense and Sensitivity. *Economic Systems Research*, 10, 307–323.
- Dietzenbacher, E. and B. Los (2000) Structural Decomposition Analyses with Dependent Determinants. *Economic Systems Research*, 12, 497–514.
- Dietzenbacher, E. and J. Stage (2006) Mixing Oil and Water? Using Hybrid Input–Output Tables in a Structural Decomposition Analysis. *Economic Systems Research*, 18, 85–95.

- Dietzenbacher, E., B. Los, R. Stehrer, M. Timmer and G. de Vries (2013) The Construction of World Input–Output Tables in the Wiod Project. *Economic Systems Research*, 25, 71–98.
- Erumban, A.A., R. Gouma, B. Los, R. Stehrer, U. Temurshoev, M. Timmer and G. de Vries (2011, June 13–17) World Input–Output Database (WIOD): Construction, Challenges and Applications (Paper presented at the 19th International Input–Output Conference, Alexandria, USA, Retrieved June 13-17, 2011).
- Eurostat (2014) Physical Energy Flow Accounts (PEFA) – Manual. Luxemburg, Eurostat.
- Foran, B., M. Lenzen, C. Dey and M. Bilek (2006) Integrating Sustainable Chain Management with Triple Bottom Line Accounting. *Ecological Economics*, 52, 143–157.
- Genty, A., I. Arto and F. Neuwahl (2012) Final Database of Environmental Satellite Accounts: Technical Report on Their Communication. WIOD Deliverable 4.6.
- Gui, S., H. Mu and N. Li (2014) Analysis of Impact Factors on China's CO₂ Emissions from the View of Supply Chain Paths. *Energy*, 74, 405–416.
- Hoekstra, R. and J.C.J.M. van der Bergh (2002) Structural Decomposition Analysis of Physical Flows in the Economy. *Environmental and Resource Economics*, 23, 357–378.
- Lenzen, M. (2000) Errors in Conventional and Input–Output-Based Life-Cycle Inventories. *Journal of Industrial Ecology*, 4, 127–148.
- Lenzen, M. (2003) Environmentally Important Paths, Linkages and Key Sectors in the Australian Economy. *Structural Change and Economic Dynamics*, 14, 1–34.
- Lenzen, M. (2007) Structural Path Analysis of Ecosystem Networks. *Ecological Modelling*, 200, 334–342.
- Lenzen, M. and R.H. Crawford (2009) The Path Exchange Method for Hybrid LCA. *Environmental Science & Technology*, 43, 8251–8256.
- Lenzen, M., R. Wood and T. Wiedmann (2010) Uncertainty Analysis for Multi-Region Input–Output Models – A Case Study of the UK's Carbon Footprint. *Economic Systems Research*, 22, 43–63.
- Lenzen, M., K. Kanemoto, D. Moran and A. Geschke (2012) Mapping the Structure of the World Economy. *Environmental Science & Technology*, 46, 8374–8381.
- Lenzen, M., D. Moran, K. Kanemoto, B. Foran, L. Lobefaro and A. Geschke (2012) International Trade Drives Biodiversity Threats in Developing Nations. *Nature*, 486, 109–112.
- Lenzen, M., D. Moran, K. Kanemoto and A. Geschke (2013) Building Eora: A Global Multi-Region Input–Output Database at High Country and Sector Resolution. *Economic Systems Research*, 25, 20–49.
- McDougall, R. and H.-L. Lee (2006) Chapter 17: An Energy Data Base for GtAP. In: B.V. Dimaranan (ed.), *Global Trade, Assistance, and Production: The GTAP 6 Data Base*. West Lafayette IN: Purdue University, Centre for Global Trade Analysis, 1–34.
- Miller, R.E. and P.D. Blair (2009) *Input–Output Analysis: Foundations and Extensions*. Cambridge UK, Cambridge University Press.
- Minx, J., G. Baiocchi, G.P. Peters, C.L. Weber, D. Guan and K. Hubacek (2011) A “Carbonizing Dragon”: China's Fast Growing CO₂ Emissions Revisited. *Environmental Science & Technology*, 45, 9155–9153.
- Moran, D. and R. Wood (2014) Convergence Between the Eora, WIOD, EXIOBASE, and OpenEU's Consumption-Based Carbon Accounts. *Economic Systems Research*, 26, 245–261.
- Oshita, Y. (2012) Identifying Critical Supply Chain Paths That Drive Changes in CO₂ Emissions. *Energy Economics*, 34, 1041–1050.
- Owen, A. (2015) Techniques for Evaluating the Differences in Consumption-Based Accounts: A Comparative Evaluation of Eora, GTAP and WIOD (Unpublished PhD thesis, University of Leeds, Leeds UK).
- Owen, A., K. Steen-Olsen, J. Barrett, T. Wiedmann and M. Lenzen (2014) A Structural Decomposition Approach to Comparing MRIO Databases. *Economic Systems Research*, 26, 262–283.
- Peters, G.P. and E.G. Hertwich (2006) Structural Analysis of International Trade: Environmental Impacts of Norway. *Economic Systems Research*, 18, 155–181.
- Peters, G.P., R.M. Andrew and J. Lennox (2011) Constructing an Environmentally-Extended Multi-Regional Input–Output Table Using the GTAP Database. *Economic Systems Research*, 23, 131–152.

- Peters, G.P., S.J. Davis and R.M. Andrew (2012) A Synthesis of Carbon in International Trade. *Biogeosciences Discussions*, 9, 3949–4023.
- Sonis, M., G.J.D. Hewings and S. Sulistyowati (1997) Block Structural Path Analysis: Applications to Structural Changes in the Indonesian Economy. *Economic Systems Research*, 9, 265–280.
- Steen-Olsen, K., A. Owen, E.G. Hertwich and M. Lenzen (2014) Effects of Sector Aggregation on CO₂ Multipliers in Multiregional Input–Output Analyses. *Economic Systems Research*, 26, 284–302.
- Strømman, A.H. (2009) Dealing with Double-Counting in Tiered Hybrid Life-Cycle Inventories: A few Comments – Response. *Journal of Cleaner Production*, 17, 1607–1609.
- Strømman, A.H., G.P. Peters and E.G. Hertwich (2009) Approaches to Correct for Double Counting in Tiered Hybrid Life Cycle Inventories. *Journal of Cleaner Production*, 17, 248–254.
- Suh, S. (2005) Theory of Materials and Energy Flow Analysis in Ecology and Economics. *Ecological Modelling*, 189, 251–269.
- Sun, J.W. (1998) Changes in Energy Consumption and Energy Intensity: A Complete Decomposition Model. *Energy Economics*, 20, 85–100.
- Treloar, G.J. (1997) Extracting Embodied Energy Paths From Input–Output Tables: Towards an Input–Output-Based Hybrid Energy Analysis Method. *Economic Systems Research*, 9, 375–391.
- Tukker, A., A. de Koning, R. Wood, T.R. Hawkins, S. Lutter, J. Acosta, . . . J. Kuenen (2013) Exiopol – Development and Illustrative Analyses of a Detailed Global MR EE SUT/IOT. *Economic Systems Research*, 25, 50–70.
- Tukker, A., E. Poliakov, R. Heijungs, T.R. Hawkins, F. Neuwahl, J.M. Rueda-Cantuche, . . . M. Bouwmeester (2009) Towards a Global Multi-Regional Environmentally Extended Input–Output Database. *Ecological Economics*, 68, 1928–1937.
- UNDESA. (2015) System of Environmental-Economic Accounting for Energy. Rome, UNDESA.
- UNFCCC. (2007) CO₂ Emissions in 2007. Retrieved April 9, 2015, from http://unstats.un.org/unsd/environment/air_co2_emissions.htm
- Usubiaga, A. and J. Acosta-Fernández (2015) Carbon Emission Accounting in MRIO Models: the Territory vs. the Residence Principle. *Economic Systems Research*.
- Wiedmann, T., H.C. Wilting, M. Lenzen, S. Lutter and V. Palm (2011) Quo Vadis MRIO? Methodological, Data and Institutional Requirements for Multi-Region Input–Output Analysis. *Ecological Economics*, 70, 1937–1945.
- Wood, R. and M. Lenzen (2003) An Application of a Modified Ecological Footprint Method and Structural Path Analysis in a Comparative Institutional Study. *Local Environment*, 8, 365–386.
- Wood, R. and M. Lenzen (2009) Structural Path Decomposition. *Energy Economics*, 31, 335–341.
- Wood, R., T.R. Hawkins, E.G. Hertwich and A. Tukker (2014) Harmonising National Input–Output Tables for Consumption-Based Accounting – Experiences from Exiopol. *Economic Systems Research*, 26, 387–409.