Chapter 2
Literature Review

Abstract This book identifies techniques that can be used to evaluate the differences in consumption-based accounts (CBA) calculated by three multiregional input–output (MRIO) databases. This literature review chapter gives an overview of the development of environmentally-extended input–output analysis, followed by descriptions of how to construct an MRIO database, reviews of the metadata documents from existing MRIO databases, summaries of studies that aim to understand uncertainty in both IO and MRIO results and descriptions of research and techniques that this book will use to understand difference in MRIO databases. The following chapter on methodology and data then focuses on the specific techniques that are employed in this study. Any mathematical equations are to be found in the methods chapter (Chap. 3).

2.1 A Brief Overview of Input–Output Techniques

Input–output analysis uses an analytical framework to describe the economy of a region, nation or even the entire world (Miller and Blair 2009). The basic framework is shown in Fig. 2.1. \( Z \) is a matrix showing inter-industry transactions; \( Y \) is final demand sales to households, government and capital investments; \( h \) is the value added in compensation of employees, taxes on production and imports less subsidies, and gross operating surplus; \( x \) is the sum of all outputs; and \( f \) is extension data such as for example pollutants, energy use or number of employees by industrial sector.

Figure 2.1 shows a symmetric IO table (SIOT), where each industry is the producer of a single product type. The transaction matrix can take one of two forms: either a product-by-product (P-by-P) IO table or an industry-by-industry (I-by-I) IO table. A P-by-P table describes the quantity of product used to make each product irrespective of the producing industry, whereas an I-by-I table describes inter-industry relations (Eurostat 2008). In reality, some industries produce two or more product types. For example Yamaha produces and sells both motorcycles and pianos! To understand instances of co-production, sometimes IO tables are constructed in a supply and use table (SUT) format as shown in Fig. 2.2. Here the \( Z \) matrix of inter-industry transactions is separated into two separate accounts; the supply matrix and the use
matrix. In the SUT system, the supply table is transposed to form the make table and denoted by $V$. It shows the products that are made by each industry. The use matrix $U$, shows the intermediate products that are bought by each industry in order to make their final products. The greyed out sections contain zeros. In the SUT format final demand is only recorded for products, and value added and environmental extensions are only recorded for industries.
In both SIOT and SUT formats, in order to understand the role demand plays in the production of goods and services, a series of linear equations are formed that describe how producing a single unit of final demand requires inputs from all sectors of the economy. Solving this series of equations reveals the production recipe required to make the product. For details of the equations see Sect. 3.1. It is generally accepted that the economist Wassily Leontief (Bjerkholt and Kurz 2006) was the sole instigator of this field and the inverse function used to solve the series of equations—the Leontief inverse—takes his name.

The discipline of input–output analysis has developed significantly since its conception in the 1930s by Leontief. Expanding from Leontief’s (1936) 41 sector model of the American economy, today’s IO analysts have the choice of several databases containing time-series data on thousands of sectors, from countries spanning the globe. The expansion and development of IO analysis has been driven by three main factors. Firstly, it has become a requirement for many countries to produce annual consistent systems of national accounts (SNA) to calculate gross domestic product (GDP). The EU (European Union) member states are required to produce standardised 60 sector SUTs on an annual basis to comply with the European System of Accounts (ESA) 95, from which a set of SIOTs are generated every five years (Tukker et al. 2009). Secondly, advances in high performance computing have meant that working with and storing very large input–output databases has become more manageable (Wiedmann et al. 2011). Finally, the political concerns of the time have influenced the type of research question IO analysis is used for. For example, in recent years, growing concern about harmful concentrations of GHGs in the atmosphere has prompted renewed interest in using environmentally-extended IO (EEIO) techniques to understand the role of demand in increases in emissions and the development of consumption-based accounts (CBA) to complement the existing territorial emissions inventories (Barrett et al. 2013; Davis and Caldeira 2010; Hertwich and Peters 2009; Minx et al. 2009a, b; Peters et al. 2011b; Peters 2008; Wiedmann and Barrett 2013).

It is impossible to say which of these three factors has been most influential in the constantly evolving IO methodology. When taking a chronological approach to reviewing IO methods and applications, one has to bear in mind the stage each of the above factors had reached when the research was conducted. For example, Leontief’s (1936) initial study built a single region IO table of the American economy to understand the effect of a change in demand on the types of jobs needed after the American recession. In the 1930s, Leontief would have had to manually solve the series of simultaneous equations to construct the Leontief inverse used in his calculations. This time consuming exercise limits the total manageable matrix size. Leontief’s (1936) IO table from 1919 included a column of American produced goods that are removed for exports and a row of imports showing inputs to the intermediate demands of US industry and a total import to final demand. These initial IO studies tended to have a single country focus and, as described, had relatively simplistic methods for dealing with traded goods.

Section 2.1.2 explains how IO analysis has evolved to take into account imports from multiple trade regions and to start to map the complex web of transactions that make up product supply chains. To explain the complexity of global trade systems
this review uses the example of EEIO analysis—the history of which is described in Sect. 2.1.1.

2.1.1 Environmentally-Extended Input–Output Analysis

Since the late 1960s researchers have theorised about accounting for externalities such as waste, production losses, scrap and pollution in production processes (Ayres and Kneese 1969; Kagawa 2012). As early as 1966, Cumberland (1966) proposed that IO techniques could be a useful methodology in understanding the consequences of development processes on the environment. These early investigations involved the inclusion of a vector of ‘externalities’ which measured tonnes of pollution per unit of output for each industrial sector. Calculations could then determine how pollution originating from producing sectors could be reallocated to final users. These early studies were static, ex-post analyses describing the situation as observed at the end of the time frame of measurement. Kagawa (2012, p. 4) explains that these types of analyses can be criticised for not considering “the abatement activities of various pollutants generated by production activities”.

In 1970, in his paper presented to the International Symposium on Environmental Disruption in the Modern World, Leontief demonstrates how an IO framework can be extended further to consider pollution abatement activities by introducing the concept of an anti-pollution industry into the inter-industry flow matrix (Leontief 1970). These types of analyses were able to estimate both the economic and polluting effects of a new government spending program. Using generalised IO methods allows researchers to optimise one or more objective functions. For example, Miller and Blair (2009) demonstrate using IO methods to minimise pollution whilst meeting a set level of final demand. This aspect of environmental IO analysis has fed into the research areas of general equilibrium modelling and macro-economic techniques. These dynamic systems are useful for future projections and policy simulations, but are outside of the scope of this book which concentrates on the comparison of static databases.

More recently, researchers have returned to focus on the information that can be gleaned from the static ex-post approach described earlier. Following the 1997 United Nations Climate Change Conference in Kyoto—and the resulting protocol whereby the world’s developed nations agreed to greenhouse gas emissions reduction targets—understanding the cause of carbon emissions has become a research priority. IO analysis can be used to gain a further understanding of the role consumption has to play in the generation of emissions (Hertwich and Peters 2009; Peters and Hertwich 2008a, b; Wiedmann et al. 2007). Using IO techniques, analysts can reassign the CO₂ emissions associated with production activities to the final demand of products. Summing the emissions associated with a nation’s demand for products, along with the direct emissions from the heating of homes and private transportation,¹ calculates

¹Known as direct household emissions.
what has come to be known as a ‘carbon footprint’ (Wiedmann and Minx 2008). The interest in EEIO further increased when researchers started to calculate and compare consumption and production emissions at a national level (Hertwich and Peters 2009; Peters and Solli 2010; Weber and Matthews 2007; Wiedmann et al. 2010). These types of calculations require information on not only the interactions between domestic industries and their associated environmental impacts, but also on what products are imported into the country, what their environment impacts are, and what domestic products leave the country as exports. To undertake this type of calculation, the IO table must accurately describe trade in some detail.

2.1.2 Understanding Trade in Input–Output Analysis

Adding a geographic extension to the basic IO framework can help understand impacts associated with trade. To consider the impacts associated with global production systems, the IO structure should to take into account impacts of production elsewhere in the world and understand how goods and services are traded globally. There are two types of flows of traded goods for which the additional impacts can be measured. Either a consumer in country A buys an imported finished good as a final demand product, or an industry in country A imports goods from the rest of the world as an intermediate demand that is then used to produce its final product. Similarly, products can leave Country A either as finished goods that are imported to other countries as final demand or as intermediate demands to other countries industry.

Figure 2.3 shows the development of how IO tables have dealt with trade as the databases themselves have increased in complexity. The single region treats each country in isolation; bidirectional trade considers how country 1 (C1) imports from and exports to each other region; and multidirectional trade understands the trade between, for example, countries 4 and 5 that contributes to products imported by country 1.

![Fig. 2.3 Development of understanding trade in IO analysis (adapted from Lenzen et al. 2004)](image-url)
2.1.2.1 Single Region Input–Output Models

In the very first IO analysis Leontief (1936) uses a single region input–output (SRIO) table. To calculate a country’s CBA it is assumed that products that are imported to intermediate or final demand are produced with the same production recipe as domestic goods and services. This is known as the domestic technology assumption (DTA).

The SRIO framework, as shown in Fig. 2.4, splits final demand into those products bought by country A’s consumer and those that are exported to other nations. This allows the analyst to understand the role domestic demand has on production. Sales to Country A final demand does not distinguish between final demand of domestic or imported products here. To complete an environmental-impact study using a SRIO database, the imports row is also used. The analyst adds the impact of intermediate imports to the account.

Despite criticism of this approach (Andrew et al. 2009; Peters et al. 2011b), SRIO based analyses were still commonly used for environmental-impact studies as recently as 2009. In a recent review of consumption-based accounting approaches using IO methods, Wiedmann (2009) cites 31 such studies published between 2006 and 2009.

Fig. 2.4 SRIO Framework
2.1.2.2 Bidirectional Trade Input–Output Models

This method uses each region’s SRIO table alongside bilateral trade data (BTD) to measure the emissions embodied in bilateral trade (EEBT). EEBT uses domestic technologies to calculate impact of both domestic products consumed domestically and the impact of those domestic products that are exported abroad both as final demands and intermediate demands to foreign industry (Peters and Solli 2010). The impact of imported goods is then calculated as the sum of every other country’s emissions embodied in their exports to the initial country. By starting with the territorial emissions in a country or region, and subtracting the balance of emissions embodied in trade (BEET), the end result is a calculation of a trade-adjusted emissions inventory (TAEI) (Peters and Solli 2010).

Figure 2.5 shows the TAEI flows for country A. The purple arrows represent final demand impacts due to country A’s consumption; blue arrows show intermediate imports to country A’s industry; green arrows show intermediate exports to other countries’ industry and red arrows show exports to other countries final demand. Country A’s TAEI is found by taking domestic production emissions and adding the purple and blue flows that flow in to the boundary and subtracting the green and red flows that flow out. Note that the boundary, (dashed oval) within which the emissions are measured, includes both the consumers in country A, and the industries.

The IO structure required for bidirectional trade IO databases is shown in Fig. 2.6. The greyed-out sections contain zeros. Here, the final demand vector represents final demand for domestic products. Bilateral trade data (BTD) distinguishes the destination country of an exporting country’s exports. The exports include both exports to final and intermediate demand. To calculate country A’s consumption based account, the ‘country A final demand’ vector and the ‘exports from countries B & C to country A’ are used.

Zhou and Kojima (2009) state that if exports of intermediate demand are treated exogenously, as in EEBT approaches, the impacts associated with the use of intermediate commodities by downstream production are not accounted for properly. In other words, the emissions associated with a textile product from China, which is bought in the UK, might contain some emissions in the supply chain that were generated in the UK as part of its production which do not get accounted for. Rather than dismiss this approach as not handling flows correctly, both Peters and Solli (2010) and Kanemoto et al. (2012) urge that practitioners need to ensure that the correct question is being asked of the model and the results are interpreted appropriately. The EEBT approach produces measures of exports and imports that are consistent with reported bilateral flows and can reveal the sizes of both final and intermediate demand. This technique can help provide an answer to the research question ‘what are the territorial based emissions in country C to produce goods and services which are imported?’ (Peters and Solli 2010).
2.1.2.3 Multidirectional Trade Input–Output Models

In a multidirectional trade model, rather than linking together separate SRIO tables using BTD, a multiregional input–output (MRIO) table is constructed. An MRIO table can be considered as one very large IO table. In the MRIO table, each column shows the industry requirements from both domestic and foreign sectors to produce a product from a specific sector in a specific country. This means that if a consumer
2.1 A Brief Overview of Input–Output Techniques

in country A, buys a domestically produced product, it takes into account any inter-
mEDIATE flows from countries B and C that are used to make products in country A
that are consumed by country A consumers. Figure 2.7 shows this as the arrows with
solid lines. Note that the purple and green solid arrows represent goods purchased
from domestic production in country A but originate from industries in countries B
and C with some processing in A. This effect is shown by the arrow passing through
country A’s industry. Also note that a product imported to final demand from country
B (dotted arrows) can include not only emissions from industry in countries B and
C, but also some domestic territorial emissions from country A.

Here the boundary is drawn around Country A’s consumers and does not include
country A’s industry. If the boundary included industry, the red arrows would be
double counted. The MRIO system can show the consumption account for country
A broken down by the country of final assembly (or the place shown in the final
demand imports) by summing the solid arrows (for country A), the dotted arrows
(for country B), and the dashed arrows (for country C). Or, alternatively, the system
can show the consumption account broken down by source country by summing the
red arrows (for country A), the blue arrows (for country B) and the green arrows (for
country C).

The IO structure required for a multidirectional trade IO database—an MRIO
database—is shown in Fig. 2.8. To calculate country A’s consumption based account,
only the final demand of country A’s consumers is used. If the MRIO framework
(Fig. 2.8) is compared with the EEBT framework (Fig. 2.6) and the SRIO framework

---

**Fig. 2.6** EEBT Framework
(Fig. 2.4), we see that additional information is needed, beyond that which is provided in each country’s SRIO. BTD informs where exports go by product type, destination region and whether this is to final or intermediate demand. This means that the final demand vector in the MRIO can be extended to show Country A’s final demand of B & C as this is the exports from B & C to country A’s final demand. However, for the intermediate demand, the framework requires not only the product type and destination, but the industry that is buying it. This means that the column ‘exports...
from Countries B & C to Country A’ in the EEBT framework has to populate the matrix of ‘imports to country A intermediate demand from countries B & C’. Since this data is missing from BTDs, MRIO databases often require some estimation in their content. In Sect. 2.2, the construction estimations are discussed in more detail.

A full MRIO database can isolate, capture and measure each of the explicit flows from every industry, in every country making up the full supply chain of a product (Su and Ang 2011; Wiedmann et al. 2011; Wiedmann 2009). Tukker et al. (2009, p. 1931) state MRIO as the “best way of taking trade into account” but again, Peters and Solli (2010) explain that this is very much dependent on the research question. The EEBT approach is the only way to count the exact size of the flows that leave a country as exports (regardless of if they flow back in imported goods). Both EEBT and MRIO account for the same global emissions but the allocation is different depending on the level of trade in intermediate products. MRIO endogenises the intermediate demand, and so the system only calculates using final demand to avoid any double counting of intermediate consumption. Because EEBT does not consider flows from B to C in A’s account, a TAEI does not double count intermediate demand either.

**2.2 MRIO Construction**

An MRIO table for $n$ countries each with $m$ sectors is a matrix of dimensions $mn$ rows by $mn$ columns and rather than considering a single nation’s economy it treats the entire global economy as a single system. As Fig. 2.8 shows, the MRIO table is constructed by placing the SRIO tables from every region along the diagonal of a large composite matrix and filling in the off diagonal matrices to show the sectoral
requirements from non-domestic regions in the production of domestic products (Peters et al. 2011a). Construction assumes that SRIO tables are available for all nations, that there is a degree of harmonisation in sectors in each SRIO and that trade linked data can be determined (Tukker et al. 2009). One of the reasons the EEBT technique has been used to account for emissions from consumption rather than a full MRIO analysis is the difficulties in obtaining suitable data to construct a MRIO table (Peters et al. 2011a). Sectors rarely match between different countries’ SNAs and populating the off diagonal sections is complex, time consuming and can involve a lot of assumptions. As Dietzenbacher et al. (2013, p. 73) state, “Constructing a large database [like in the WIOD project] implies that several choices need to be made”. In Sects. 2.2.1 and 2.2.2 the data requirements and data manipulations needed to construct an MRIO are discussed in detail.

2.2.1 Data Requirements to Extend IO to Consider Global Trade

An MRIO database requires a set of SRIO tables for each country in the world and further additional data to understand the complex web of international trade interactions that take place between each country. As mentioned in Sect. 2.1, the EU member states are required to produce standardized 60 sector SUTs on an annual basis to comply with ESA95, from which a set of SIOTs are generated every five years (Tukker et al. 2009). Other major nations produce SUTs and SIOTs but there is no global standardisation to sector classification (Tukker et al. 2009). In the construction of country level SRIOs a domestic table is produced alongside either an imports row, or an imports table. An imports table is not broken down by country, so the tables show the product that is imported and the importing industry, but not the country it is imported from. An imports row simply shows the spend on imports required by each industry to produce their product. As explained in Sect. 2.1.2.2 bilateral trade databases (BTD) provide information on exports and imports of goods, broken down by trading partner country or region and the economic activity described—whether the flow is to final or intermediate demand OECD (2014a). BTDs show the amount of goods by sector that flow to and from every world region. The destination is recorded as final demand or intermediate demand to a country but for intermediate demand, it is not specified which sector destination the flow is to.

In addition to information describing the economic interactions in global supply chains, emissions data by global production sectors is required as an input for an EE-MRIO. For the EU member states’ 60 sector SUTs and SIOTs, matching sector emissions data is available from the National Accounting Matrix including Environmental Accounts (NAMEA) (de Haan and Keuning 1996). For a global system, consistently produced emissions data is needed for every country in the database. Two approaches can be used to assign an environmental impact to each industrial sector. The International Energy Agency (IEA) produce tables showing energy out-
2.2 MRIO Construction

put by activity by country and authors such as Shimoda et al. (2008) explain how emissions are matched to this data. However this ‘top down’ technique is criticised by Tukker et al. (2009) who remind us that not all countries are signatories of the IPCC (Intergovernmental Panel on Climate Change) so do not have to report such statistics. An alternative method involves estimating the CO₂ emissions associated with an industry based on the reported energy use of each sector. To do this, emission factors are applied to the energy use by industry. However, this ‘bottom up’ technique incurs the problem of global emissions totals not summing to the reported global totals Tukker et al. (2009).

2.2.2 Preparation of Data for MRIO

Before the SRIO tables and the BTD data can be combined together to produce an MRIO table, a harmonisation procedure is often required. If there are different sectoral classification systems used in the SRIO tables and BTD, a process of aggregation and disaggregation might be necessary to produce a single common classification for all nations. In addition, there are a number of conditions that the system needs to satisfy in order for the allocation functions to work: namely, the inputs to the system need to be equal to the outputs. In a global trade perspective, this means that reported imports of commodity x from country A to country B needs to be the same as the reported exports of commodity x from country A to country B. This phenomenon, known as a “mirror statistic” rarely hold true and MRIO databases need to go through balancing procedure.

Even if SRIO tables and BTD are available for each and every global region, there is still considerable work required in constructing a fully functioning MRIO table. Inomata et al. (2006), in their papers to accompany the Asian international input–output table (AIOT), describe three stages of pre-preparation before data is subjected to the balancing procedures necessary for MRIO conditions to be met.

2.2.2.1 Adjustment of the Presentation Format

The first phase—adjustment of the presentation format—involves identifying that each country’s system of national accounts reflects the differing situation within each country as to how data is collected and what is available (Inomata et al. 2006). An MRIO table needs to be consistent in the meaning of each category so that the system is comparable and can work together as a whole. Most obviously, this means that each SRIO table needs to be in the same currency. Exchange rates can be used to convert data to one common currency (Bouwmeester and Oosterhaven 2007). Additional changes that might be required to adjust the presentation of the national SRIO tables used in an MRIO table include converting data from basic prices to producer prices; adjusting the import matrices so that they are valued at CIF (cost, insurance and freight) and that they do not include import duties and commodity taxes; and dealing
with negative entries, representing government subsidies, by treating the entity as value added items. For more detail see Inomata et al. (2006). The authors recognise that there are no hard and fast rules to this procedure and there are trade-offs between a consistent and uniform system and level of original information and detail (Inomata et al. 2006).

In addition to adjustment of the economic data, the supplementary data such as kilotonnes of emissions, thousands of employees or volume of water by industrial sector must also have the same meaning. In the case of emissions, MRIO database compilers must decide whether the residence or territorial principle is applied. 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 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. The territorial principle allocates emissions to the country where they take place and are used in national statistics. This decision specifically affects how total global emissions are distributed between industrial emissions (f in Fig. 2.4) and those emissions directly from households. Emissions associated with transportation industrial sectors are also affected.

2.2.2.2 Preparation of Sector Concordance and Supplementary Data

Once data in the SRIOS have the same meaning across all tables, each table then has to be aggregated or disaggregated to a common set of sectors. Inomata et al. (2006) call this stage preparation of sector concordance and supplementary data. Each national economy has its own unique characteristics and the sector classification system used to record data reflects this character. Some economies are heavily agriculture-based and these countries will often use sector classification systems that are very detailed in the agriculture sectors, whilst other might be more biased to industry. An additional consideration is the total number of sectors recorded, Inomata et al. (2006) aggregated the 517 sectors for Japan to their consistent set of 76 sectors for the AIIOT system. Bouwmeester and Oosterhaven (2007) note that often it is easier to revert to older classification systems when attempting to produce a common set of sectors. Summing two or more sectors to a single new sector is a simple enough procedure. Inomata et al. (2006) note that the difficulties that arise when a national IO entry needs to be split between two or more sectors in the new consistent sector system because additional data is needed to do this. Alongside a consistent set of SRIOS, the BTD and the additional industry supplementary data must also map to the consistent set of sectors.

Sets of SRIOS tables do not cover every county in the world. For an MRIO to function without losing information, a ‘rest of world’ (RoW) region is often required to describe the trade flows of countries that have not produced SRIOS tables. The volume of trade by sector and country can be estimated by looking at the differences between reported global trade flows and the sum of flows by countries whose data
has been captured. The missing element is a generalised structure of the economy for the RoW—a RoW SRIO. One approach is to pick a country that is considered representative of the RoW (Peters 2007a). The selection of this representative country will depend on which countries there are already data for. For example, some authors studying specific continents, such as Europe might choose China’s SRIO to represent the RoW (Peters 2007a). Nakano et al. (2009), when using the OECD SRIO tables to calculate EEBT, used the emissions factors of Malaysia to represent the RoW. For their work on the AIIOT MRIO, Su and Ang (2011) argue that the RoW region behaves similarly to the average Asian economy, noting similarities in the per capita GDP of the RoW and Asia and the emissions intensities. The authors aggregated nine Asian economies to simulate the emissions intensities and domestic SRIO table for the RoW. The final demand structure was also mirrored for RoW final demand (Su and Ang 2011).

2.2.2.3 Reconciliation of Data and Balancing the Table

The SRIO tables, modified to common currencies and sectors, are then placed in an MRIO table. The final stage is *reconciliation of the data and balancing the table* (Inomata et al. 2006). The first stage in the balancing procedure is setting up the off-diagonal matrices of the MRIO. Consider a set of $n$ regions and $m$ sectors in an MRIO system. Region $k$, will sell to and buy from ‘$n - 1$’ other regions. This means that within the column representing who region $k$’s $m$ industrial sectors buy from, a stack of ‘$n - 1$’ additional trade matrices is needed along with region $k$’s SRIO table. Import tables reveal how much each industrial sector imports and sometimes they distinguish which products are imported (Tukker et al. 2009). However, the import tables do not reveal the country of origin, i.e. which of the $n - 1$ regions the import flow is from. These, import tables can be disaggregated to show region of origin using BTD (Bouwmeester and Oosterhaven 2007). However, BTD gives detail on the product that is being imported, where it is being imported from, which country is importing it, but not which industrial sector it is destined to be used for. Clearly assumptions have to be made to fill in the missing parts of the puzzle and there are a number of methods that can be used. Sections 2.3.1–2.3.3 explain how GTAP, WIOD and Eora respectively deal with this issue.

Inomata et al. (2006) describe the table, at this stage, as being balanced with respect to input composition, but that demand and supply for each country are not consistent. The sum of flows of particular sector from a particular country to all countries of destination should equal the reported export by that country of origin in the BTD, however as Tukker et al. (2009), Inomata et al. (2006) and Bouwmeester and Oosterhaven (2007) note, this is rarely the case. Inconsistencies occur due to differences in sector classification systems, exports being wrongly assigned to countries that goods pass through the ports of rather than the actual country of origin and other reasons that will be discussed fuller (Peters et al. 2011a; Andrew and Peters 2013).
The table then needs to be balanced, and this is often done using a method known as RAS. The RAS technique uses an iterative process to alter individual cell values using the known export columns and import rows of the original IO tables as constraints (Bouwmeester and Oosterhaven 2007). Because the domestic SRIO tables are treated as known data, before applying the RAS technique to the MRIO, sometimes these tables are removed and replaced with zeros. One of the consequences of the RAS procedure is that it will re-price the import matrices from CIF to be in FOB (Free On Board) matching the export prices.

### 2.3 Data Sources and Construction of Current MRIO Systems

The latest audits of the main global MRIO initiatives (Inomata and Owen 2014; Peters et al. 2011a; Wiedmann et al. 2011), describe six MRIO databases of which four were launched in or after 2012 (see Table 2.1) although there is concern that some systems may not be updated regularly due to funding dependencies (Peters et al. 2011a). This study chose to compare CBA for the year 2007 because, at the time of writing, it is the latest year where there are at least three EE-MRIO databases to compare. The three MRIO databases chosen are Eora, GTAP and WIOD. The literature review continues by assessing the metadata and construction techniques specific to these three MRIO databases. The review starts with GTAP since the database has been in existence for the longest time and the construction method is the simplest. WIOD is reviewed second and Eora last because this database differs most in construction methodology. Finally, Sect. 2.3.4 compares the three MRIO databases chosen for this study.

#### 2.3.1 GTAP MRIO

The Global Trade Analysis Project is described as “a global network of researchers and policy makers conducting quantitative analysis of international policy issues”. GTAP’s goal is to “improve the quality of quantitative analysis of global economic issues within an economy-wide framework” (GTAP 2014a). GTAP was not initially designed as an MRIO database and is mainly known for its use in CGE modelling (GTAP 2014b). Since the project provides tables of intermediate demand, final demand, bilateral trade and an emissions extension, researchers looking to construct MRIO databases, turned to GTAP. Presenting at the 16th International input–output Association (IIOA) conference, Peters (2007b) first suggested the suitability of the GTAP data for use in constructing an MRIO database and later demonstrated how it could be used for global MRIO studies (Hertwich and Peters 2009; Peters and

---

2 January 2015. EXIOBASE was not freely available at this time.
Table 2.1 MRIO systems currently available

<table>
<thead>
<tr>
<th>MRIO</th>
<th>Region detail</th>
<th>Sector detail</th>
<th>Time series</th>
<th>Extensions</th>
<th>Status (as of January 2015)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eora</td>
<td>188</td>
<td>Vaires by country, ranging from 26 to 511</td>
<td>1970–2013</td>
<td>Energy, emissions, water and land footprints, employment</td>
<td>Released in 2012. Updated annually</td>
</tr>
</tbody>
</table>

Hertwich (2008a). The advantages of using an MRIO, in this case one built from GTAP v6 data, rather than the using the domestic technology assumption (DTA) is explored by Andrew et al. (2009). In 2011, Peters et al. (2011a) published the full details of how to construct an MRIO from the GTAP v7 database.

2.3.1.1 The Original Database

The data in the GTAP database is sourced from voluntary submissions from GTAP users rather than being data taken directly from national statistical offices (Walmsley

---

3 GTAP version 6 has 87 regions and 57 sectors.
4 GTAP version 7 has 113 regions and 57 sectors.
The submissions have to meet a set of criteria and checks, such as having a minimum number of sectors; being balanced; and having an IO structure similar to an average IO table Walmsley and Lakatos (2008, p. 3). Peters (2007b) criticises the source data by claiming that it is often not up to date and, in the same release, data from different years for different countries will be supplied under the overall claim of being a 2007 dataset. GTAP resolves this issue in the same procedure it uses for converting to a common currency. The tables are scaled to the 2007 USD value converted using Market Exchange Rates (MER). Peters (2007b) notes that this method assumes an equal rate of inflation across all sectors and that in IO databases, basic prices are preferred (Peters et al. 2011a).

In the version 7.1 GTAP database used in this study, 58 out of the total 113 regions needed some form of disaggregation to convert the tables to the 57 required product sectors. GTAP tables are in the product-by-product (P-by-P) format. For every country, the non-agricultural sectors are disaggregated using a representative table formed from the set of IO tables which have the full sectoral disaggregation (Narayanan 2014; Walmsley and Lakatos 2008). The agricultural sectors are disaggregated using an additional database built partially from FAO (Food and Agriculture Organisation) data (Peterson 2014; Walmsley and Lakatos 2008). Rather than having a single RoW region, GTAP v7.1 contains 20 composite regions such as ‘Rest of South East Asia’ which are calculated as a linear combination of the known IO tables for that region and matching the required income level for the area.

One area where GTAP does not rely on user submitted values is in the energy rows of the IO tables. Here physical data on energy use in Joules is taken from the International Energy Agency (IEA), converted to monetary values and placed in the IO tables (Peters et al. 2011a). The same IEA energy data is used to generate the CO₂ emissions extension data but GTAP uses different assumptions compared to the IEA when converting energy to CO₂ (Peters et al. 2012). Emissions from bunker fuels are dealt with differently to IEA. In addition, the GTAP CO₂ emissions only cover fuel burning emission and do not include process emissions from cement Lee (2008). GTAP uses the territorial principle for emissions allocation but allocates international transportation to consumers not producers (Peters et al. 2011a).

The Bilateral Trade Data (BTD) supplied by GTAP is sourced from UN Comtrade but undergoes a process of reconciliation from its original state. The UN Comtrade database is a collection of countries reported imports and exports by commodity. A country reports what products were imported from which countries and what products were exported to which countries. This means that the same traded good should be reported twice. For example spend on footwear imported to the UK from Italy should equal the reported export of Italian footwear to the UK. However there are discrepancies in the recorded transactions. GTAP resolves this issue by measuring the reliability of each reporting country and calculating whether a nation systematically over or under reports trade (Gehlhar 2001). When deciding which of the pair of transaction costs to choose to keep in the BTD, GTAP simply checks the reliability index of each of the country and chooses the data from the country that scores best (Gehlhar 2001). This means that the BTD supplied by GTAP is already balanced—a requirement for use in CGE modelling Peters (2007b). Peters (2007b) has some
2.3 Data Sources and Construction of Current MRIO Systems

concerns about the level of data manipulation within the GTAP data and highlights particular examples of nonsensical values that may have arisen as a result of the calibration process.

2.3.1.2 Converting to an MRIO

Peters et al. (2011a) describe in detail the process for converting the data in GTAP into an MRIO system. One of the main considerations is that—as described in Sect. 2.2.2.3—the format of BTD is a vector showing commodity and import country and for an MRIO, rather than a matrix which would include destination sector. This vector needs to be stretched across both one of the off-diagonal sections and the imports to final demand, (shown in Fig. 2.8) so needs the importing sector information to provide the horizontal dimension. Peters et al. (2011a) explain how bilateral exports are distributed according to the import structure in the importing region which ensures that the output balance is conserved. Peters et al. (2011a) argue that without the knowledge of any additional information, using the import structure as a proportional distribution is as good an assumption as any. This means that each row of the off-diagonal matrices, which represent intermediate imports, has the same proportional breakdown across destination sectors. Another limitation of this technique for disaggregating country of origin based on total global averages is that each industry \( j \) in region \( s \) buys the same percentage of products from industry \( i \) in region \( r \) (Bouwmeester and Oosterhaven 2007). In other words, if UK industries are importing steel and Mexico is the country of origin for 60% of all of the steel that is imported by the UK, then for every industry in the UK, 60% of steel imported to domestic production will always come from Mexico regardless of the destination industry. In addition, imports of steel to final demand will have the same proportion—60%—of steel from Mexico. This assumption is likely to introduce error when assessing the impacts of product from places whose domestic production is heavily reliant on imported components.

2.3.2 WIOD MRIO

The World input–output Database (WIOD) was a European Commission seventh framework programme funded project running from May 2009-April 2012 (Dietzenbacher et al. 2013; WIOD 2014). Unlike GTAP, WIOD was always designed to be used for MRIO analysis and the developers state the following initial aims for the database: it must be global; cover change over time; include a variety of socio-economic and environmental indicators; and be presented in a coherent framework (Dietzenbacher et al. 2013).

WIOD takes published national statistical agencies’ SUTs as its initial data source because, as Dietzenbacher et al. (2013) argue, the SUT better represents co-production. These national tables are harmonised to a 59 product, 35 industry
common classification using a set of concordance matrices developed for the WIOD project. Sometimes this involved disaggregation of particular industries or products using common industry or product shares. If there are missing years in a country’s set of SUTs, national accounts data is used as a constraint to update a previous years’ SUT using an SUT-RAS method (Dietzenbacher et al. 2013). Supply tables are already presented in basic prices, but the use tables, which are usually in purchasers prices, have to be converted to basic prices. The tables are also converted to USD using data from the IMF.

The next stage is to split the use tables into a table of domestic use and a table of imports, then each cell of the import use table must be split by import region (Dietzenbacher et al. 2013). To extract the imported use table from the total use table, total imports by product are taken from the supply table and the portion that is imports to final demand and investment is removed (using proportions from BTD). BTD is taken from UN Comtrade and trade in services is determined using data from the UN, Eurostat and the OECD, with the UN being the preferred source (Dietzenbacher et al. 2013). In contrast to GTAP, WIOD treats imports to intermediate demand, final demand and investments differently and allows each destination to have their own specific import share from the BTD. When Erumban et al. (2011, p. 11), explaining the construction of WIOD, state that “each cell of the import use table is split up to the country of origin where country import shares might differ across use categories, but not within these categories” by ‘use’ they means the difference between final use and intermediate use. WIOD suffers the same assumption as GTAP whereby the steel bought as intermediate demand by two different sectors have the same proportion from Mexico regardless of purchasing sector.

In contrast to GTAP, WIOD has a single RoW region. To determine the RoW imports and exports by product and country, the Global totals are found in the UN Comtrade database and the sum of the 40 WIOD countries is subtracted from this total Dietzenbacher et al. (2013). Once all the trade data is collected, RAS is used to to reconcile it. Dietzenbacher et al. (2013) point out that this procedure adjusts all the BTD from that collected at source.

The final stage is to convert the SUTs and reconciled BTD into a World SIOT. The means that the supply and use tables have to be compacted together to a single industry-by-industry table for each country. There are two methods of translating SUT into SIOTs: the fixed industry sales structure assumption or the fixed product sales structure assumption. WIOD uses the second method where, regardless of the industry producing the product, products in the supply table are reallocated according to the allocation of the industry that they would be a principle output of Dietzenbacher et al. (2013), (Eurostat, 2008). This produces an industry-by-industry table (I-by-I). A RoW intermediate use table and RoW domestic final demand block is constructed from weighted average shares from the BRICIM countries with row and column totals from UN national accounts.

---

5See Dietzenbacher et al. (2013) Table 1 for an example of this method.

6Brazil, Russia, India, China Indonesia, Mexico.
In contrast to GTAP, WIOD uses the residence principle for emissions allocation (Genty et al. 2012). For countries where emissions inventories, such as the UNFCCC inventory, are available, these datasets were matched to the WIOD sector breakdown and used as the CO₂ emissions data. If inventories were not available, emissions were estimated from the energy accounts. CO₂ emissions data is calculated by “applying CO₂ emission coefficients to emission relevant energy use and then adding process-based emissions” (Genty et al. 2012, p. 3). The countries that do not need to report to the UNFCCC, and hence are not included in its inventory but are WIOD countries are Brazil, China, South Korea, India, Indonesia, Mexico and Taiwan.

2.3.3 Eora MRIO

Eora was developed by the Integrated Sustainability Analysis (ISA) group, within the School of Physics at the University of Sydney. Lenzen et al. (2013, p. 13), describe their aims for their system as having “the maximum possible level of detail”; a time series back to 1970; minimisation of assumptions; closeness to raw data; estimates of standard deviations; and for it to be freely available for research and updated in a timely manner.

With one of Eora’s aims being to be close to raw data, where possible the SRIOs are sourced from national statistical offices. SRIOs are also taken from Eurostat, IDEJETRO and the OECD. Lenzen et al. (2013) explain that 74 national SRIOs were collected in this way. Eora also keeps the original sector classifications of the data, and maintains the SIOT or SUT format alongside keeping SIOT data in its original I-by-I or P-by-P format. This means that the Eora MRIO is not in a harmonised sector format, rather the sectors are heterogeneous and different for different countries. Thus the first few stages of data adjustment as described by Inomata et al. (2006) are skipped. For countries where there are no IO tables produced, a proxy IO table is produced. These tables combine country specific macro-econometric data with a template based on the average of the Australian, Japanese and United States tables Lenzen et al. (2012a). Bilateral trade data is sourced from UN Comtrade and UN Service trade.

The main principle behind Eora’s construction is the development of an initial estimate and the collection of raw data. An initial estimate is determined for the year 2000 and balanced and reconciled. This table becomes the initial estimate for the year 2001 and new 2001 raw data is collected and used as constraints to rebalance this table and generate a new 2001 estimate. This table can then become the starting point for 2002 and so on Lenzen et al. (2012a). Eora uses a ‘constrained optimisation algorithm’ to find a solution that best fulfils the constraints. The constraints can never be completely satisfied because it is often the case that they conflict with each other. The ISA team have developed a version of RAS called KRAS to deal with conflicting constraints (Lenzen et al. 2009). The adjustment to a common currency occurs during the optimisation routine and data from IMF is used to convert all data to US dollars (Lenzen et al. 2013). Eora is also unique in the fact that it does not
calculate a RoW region. Eora contains data from 188 countries and assumes that this covers the global economy sufficiently.

Eora does not correct for the residence principle (Lenzen et al. 2012a) and CO₂ data is sourced from the Emissions Database for Global Atmospheric Research (EDGAR) and is an initial estimate alongside data from multiple other sources such as the UNFCCC. The optimiser is then used to resolve data conflicts (Lenzen et al. 2012a).

### 2.3.4 Comparing the Source Data, Structure and Construction of Eora, GTAP and WIOD

Table 2.2 (adapted from Owen et al. 2014) provides summary information about the source data and construction techniques used in building the Eora, GTAP and WIOD MRIO databases described in Sects. 2.3.1–2.3.3. It is clear that the models differ in a number of ways. Different source data is used in both the economic and environmental extension sections of each database. GTAP uses P-by-P SIOTs, WIOD I-by-I SIOTs and Eora uses a mixture of SUTs and SIOTs. Even if the data is from the same source, each system organises it in different ways. Eora keeps the data in its original format, whereas GTAP and WIOD reorganise tables to 57 and 35 sectors, respectively. In addition, GTAP realigns energy use by sector to match the spread of joules reported by the IEA. WIOD uses the residence principle for emissions allocation whereas GTAP and Eora take the territorial approach.

Assumptions are made when data is missing and each MRIO deals with missing data in a different way. For example, WIOD constructs a single RoW region with an ‘average’ production structure, whereas GTAP models several regional RoW regions. Eora attempts to construct production structures for every national economy negating the need for a RoW region. Another element where there is missing data that needs to be constructed is in the off-diagonal trade matrices. GTAP uses a fairly blunt proportional assumption to turn a vector of import data by source into a matrix where use is the second dimension. WIOD takes care to distinguish between whether the use is intermediate or final use but the proportionality assumption remains within intermediate use sectors. Eora has a different approach recording all data on intermediate and final imports as constraints and modelling the off-diagonal matrices as a solution in the matrix optimisation process.

### 2.4 The Future of MRIO Databases

Since commencing this study a number of new MRIO systems have been developed (see Table 2.1). In this section, EXIOBASE and the OECD ICIO are briefly introduced.
### Table 2.2 Global MRIO databases used for comparisons in this study and their features

<table>
<thead>
<tr>
<th>Eora</th>
<th><strong>Source data</strong></th>
<th><strong>Availability and updates</strong></th>
<th><strong>National IO tables</strong></th>
<th><strong>Bilateral trade data</strong></th>
<th><strong>Environmental accounts</strong></th>
<th><strong>Value added data</strong></th>
<th><strong>System structure</strong></th>
<th><strong>System construction</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1970–2013 (economic data)</td>
<td>1990–2012 (extension data)</td>
<td></td>
<td>EDGAR, UNFCCC, IEA</td>
<td>National IO tables</td>
<td>188 countries</td>
<td>Uses original classification from national accounts</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Yearly updates with a 2 year lag</td>
<td></td>
<td>Territorial principle</td>
<td>UN National Accounts Main Aggregates Database</td>
<td>Varies by country; ranges from 26 to 511 sectors</td>
<td>Converts national currencies into current US$ using exchange rates from IMF</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>74 IO tables from national statistical offices</td>
<td></td>
<td>This study uses the ‘Carbon emissions from fuel burning’ account supplied by Eora</td>
<td>UN National Accounts Official Data</td>
<td></td>
<td>Off-diagonal trade data calculations, balancing and constraints</td>
</tr>
</tbody>
</table>

(continued)
|                             | Updated on a 3 year interval with a 4 year lag |
| National IO tables          | Tables submitted by GTAP consortium members |
| Bilateral trade data        | Trade in goods from UN Comtrade database |
|                             | Trade in services from UN Service trade database |
| Environmental accounts      | CO$_2$ derived from IEA energy data |
|                             | Territorial principle with reallocation of international transportation to consumers |
|                             | This study uses the data supplied by GTAP v7.1 which includes CO$_2$ from fossil fuel burning only (Lee, 2008) |
| Value added data            | Tables submitted by GTAP consortium members |
| System structure            | Region detail | 129 regions (81 for 2001) |
| Structure of IO tables      | Homogenous SIOT table structure |
| System construction         | Harmonisation of sectors | To disaggregate a country's non-agricultural sectors, the structure from other IO tables within regional groupings is used. For agricultural sectors data from the FAO is employed |
|                             | Harmonisation of prices and currency | IO tables scaled to US$ using GDP data from the World Bank |
|                             | Off-diagonal trade data calculations, balancing and constraints | BTD from UN’s COMTRADE database is harmonised, off diagonals are estimated by applying imports share across each row. No balancing required |
### Table 2.2 (continued)

<table>
<thead>
<tr>
<th>WIOD</th>
<th>Source data</th>
<th>Availability and updates</th>
<th>1995–2011 (economic)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>1995–2009 Environmental</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Funding dependent</td>
</tr>
<tr>
<td></td>
<td>National IO tables</td>
<td>SUTs from National Accounts</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bilateral trade data</td>
<td>Trade in goods from UN Comtrade database</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Trade in services from UN, Eurostat and OECD</td>
</tr>
<tr>
<td></td>
<td>Environmental accounts</td>
<td>Residence principle</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Emissions from NAMEA or estimated from energy</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>This study uses the CO₂ emissions by industry provided</td>
</tr>
<tr>
<td></td>
<td>Value added data</td>
<td>SUTs from National Accounts</td>
<td></td>
</tr>
<tr>
<td>System structure</td>
<td>Region detail</td>
<td>40 countries and a rest of the world region</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sector detail</td>
<td>35 homogeneous I-by-I sector tables</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Structure of IO tables</td>
<td>Homogenous SIOT table structure</td>
<td></td>
</tr>
<tr>
<td>System construction</td>
<td>Harmonisation of sectors</td>
<td>Developed concordance tables between national classifications and the 35 sectors used in WIOD</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Harmonisation of prices and currency</td>
<td>Supply table (from SUT) in basic prices. Use table in purchases prices.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Transform the Use table to basic prices</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Convert all data to current US$ using exchange rate from IMF</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Off diagonal trade data calculations, balancing and constraints</td>
<td>BTD finds import proportions for intermediate and final use by product. Proportions applied to import use table to split each cell by import region. International SUTs merged to a World SUT then transformed to a WIOT using the fixed product sales structure assumption</td>
<td></td>
</tr>
</tbody>
</table>
in Sects. 2.4.1 and 2.4.2, respectively. Section 2.4.3 gives an overview of the future of MRIO development.

2.4.1 EXIOBASE

EXIOBASE takes the harmonised EU SUTs as a starting point and includes more regions,\(^7\) disaggregates to 163 industrial sectors and 200 products, and combines with an extension database containing 80 resources and 40 emissions types (Tukker et al. 2013). EXIOBASE differs to GTAP and WIOD with the resulting MRIO being SUT based rather than SIOT.\(^8\) Eora, of course, is a hybrid of SIOT and SUT. After separating the imports use from the total use tables, as described in the WIOD methods (Sect. 2.3.2), and disaggregating all SUT to 129 sectors, EXIOBASE uses a nonlinear programming approach to ensure that the row and column total balance. Emissions data in EXIOBASE differs from WIOD and Eora by uses a bottom up approach by calculating from the energy using sectors. EXIOBASE calculates off diagonal trade in much the same way that WIOD does, using trade shares from UN Comtrade and UN Service data and “assuming that each industry and each final demand category imports the same share of a given product from the exporting country” (Tukker et al. 2013, p. 58). Like WIOD, EXIOBASE takes the residence principle to emissions allocation.

2.4.2 OECD ICIO

The OECD Inter-Country input–output (OECD ICIO) database is an MRIO based on national statistical agency SIOTs and SUTs. With 56 regions and 37 sectors (OECD-WTO 2012). National authorities provide data to the OECD, preferably in basic prices with both domestic and imported use tables. If this split is not provided, the OECD separates out the imports. In a joint OECD-WTO (2012) note, the issue with the proportionality assumption is highlighted. The OECD ICIO plans to explore the way imports are allocated to users but it is not yet clear how this particular MRIO has improved upon the assumption. The OECD is in the process however of developing a bilateral trade database by industry and end use category which will help improve the accuracy of the off diagonal matrices considerably. At present\(^9\) there are no environmental extensions in the OECD ICIO database but it is understood that this is something that will be considered for future development.

---

\(^7\) A total of 43 countries plus a RoW region.

\(^8\) Both I-by-I and P-by-P SIOTs are available, produced from the SUT.

\(^9\) January 2015.
2.4.3 Further Considerations

In their 2011 paper on the future directions of MRIO, Wiedmann et al. (2011) call for a number of developments within the field of MRIO research. These include hybridisation with life cycle assessment (LCA) to further improve sector disaggregation; avoiding information loss through aggregation; greater country coverage; better extension data that is relevant for sustainability research; more timely updates; historical time series; improvements in automation; transparency and testing of assumptions; and a better understanding of uncertainty. Peters et al. (2011a, p. 150) also call for “a structured comparison of the datasets to determine the necessary level of detail, accuracy and resources needed for the long-term development of environmental MRIO modelling”.

There clearly is a distinct requirement for work to carried out which understands the differences between MRIO databases and that attempts to relate the differences in outcome to the variation in source data used and the assumptions made in the database construction.

2.5 Differences in MRIO Outcomes

At the IIOA conference in Japan in 2013, a special session was arranged dedicated to exploring difference in MRIO databases. As a result of this session, a special issue of Economics Systems Research (ESR) was published in September 2014, guest edited by Anne Owen and Satoshi Inomata and included the paper which part of Chap. 5 of this book is based on (Owen et al. (2014)). While this particular paper is not discussed in the literature review, many of the examples in the following sections draw from the other studies that made up the special edition.

2.5.1 Exploring the Effect of Data and Build Choices on MRIO Outcomes

As Sects. 2.2 and 2.3 explain, there are a myriad of choices that can be made in constructing an MRIO database. Dietzenbacher et al. (2013, p. 73) explain that

these choices are often directed by the particular applications the constructors have in mind when designing the database and its underlying fundamental principles. Uncovering these is important in order to understand the differences between various alternative databases.

There have been a number of studies investigating the effect that different choices have on the outcomes produced by an MRIO and how variations in the data affect final CBAs.
2.5.1.1 Alternate Choice of Source Data

Peters et al. (2012) investigate how model outcomes change when different CO₂ emissions data are used with the GTAP MRIO. The authors investigate the effect on CBAs when emissions datasets from Carbon Dioxide Information Analysis Center (CDIAC), the UNFCCC, EDGAR, GTAP and an updated version of the GTAP data—GTAP-NAMEA are used in conjunction with the GTAP economic data. GTAP-NAMEA includes process emissions and redistributes the emissions according to the residence principle rather than the territorial technique described in Sect. 2.3.1. The study compares the average range in both production and consumption emissions for each country in the dataset and discovers that for production the average range is 30% and for consumption, 16%. Peters et al. (2012) suggest that this is because the countries that are large trade partners have lower differences in accounts. The authors also conclude that much of the difference in model outcomes “are not a reflection of the uncertainty in consumption-based estimates, but rather these differences result from the use of different production-based emissions input data and different definitions for allocating emissions to international trade” Peters et al. (2012, p. 3247).

In addition to considering the emissions data, several attempts have been made to quantify standard errors of each of the input to MRIO databases, but often these data are underreported or unavailable. For example, Lenzen et al. (2010) collect standard deviations (SD) associated with the underlying source data used to make the UK IO accounts and then regress the standard deviations across the values in the supply and use tables. This work is further explained in Sect. 2.5.1.3.

2.5.1.2 Alternate Choice of Construction Method

One method for understanding the effect of build assumptions is to build several versions of the MRIO each with different build techniques and observe the effect on the output. The types of build assumptions that can be investigated include MRIO structure and harmonisation, techniques for dealing with missing data and techniques for system balancing. Peters and Solli (2010), for example, investigate the implications of different numbers of sectors by quantifying the difference in Nordic footprints using the GTAP data first with eight aggregated sectors and then the full 57 sectors and find that the difference in CBA was relatively small. The authors state that for a “national level carbon footprint, the [MRIO database] probably does not need a high level of sector detail” (Peters and Solli 2010, p. 49). Andrew et al. (2009) perform a similar analysis on the number of countries and regions required for accurate CBA. The study finds that results can be generated that are close to those calculated using the full 113 region, but use fewer regions. However, the choice of trade regions makes a difference to the accuracy of the results.

Steen-Olsen et al. (2014) consider the sectoral breakdown in each of Eora, EXIOBASE, GTAP and WIOD and develop a common classification (CC)\textsuperscript{10} of 17

\textsuperscript{10}This classification is the one used later in this book. See Sect. 3.6 for details.
sectors which each of the MRIO databases can be mapped to. One of the features of the CC is that each sector is a one-to-one mapping with an identical sector in at least one of the full MRIO databases. Steen-Olsen et al. (2014) are then able to comment on the effect of using an aggregated multiplier because they can compare full versions of the MRIO system with its aggregated version. Interestingly, the study points out that sector aggregation does not just affect the multipliers of sectors that have been aggregated. In each of the MRIO databases, the construction sector remained a single sector in the CC but its multiplier was affected significantly by the aggregation of other sectors.

The choice of method to convert to a common currency was investigated by Weber and Matthews (2007) who show that this decision can greatly affect the size of emissions embedded in imports from certain developing countries to the USA. The authors show that choosing Purchasing Price Parity,\(^{11}\) over Market Exchange rates increases flow sizes by a factor of two for Mexico and four for China.

Stadler et al. (2014) focus on the method of constructing a Rest of World (RoW) region for MRIO databases. The authors experiment with estimating the economic structure of the RoW using every country’s SUT from the EXIOBASE database and use other various methods to determine RoW final demand resulting in 186 different RoW tables. Stadler et al. (2014) find that model runs using Switzerland and Sweden as representative RoW structures produce outlier results. Another interesting finding is that different types of CBA are affected more by the different RoW structures. For example, emissions accounts are more robust and show less variation than the land use accounts.

As described in Sect. 2.3.3, Eora’s optimisation routine for determining the off-diagonal sections of the MRIO is quite different to the approaches used by EXIOPOL, GTAP and WIOD. Geschke et al. (2014) experiment with taking the source data used for EXIOPOL and the matching constraint data used to build EXIOPOLs off diagonal trade blocks but use the Eora constraint optimisation technique (Geschke et al. 2014) to populate the off diagonal blocks. Matrix difference statistics are used to compare the original EXIOPOL table with the new version and show that there is a good correlation.

Finally, Wiebe and Lenzen (2016) explore the effect that RAS balancing techniques have on output production matrices. The Global Resource Accounting Model (GRAM) is based on OECD IO and BTD and instead of using RAS balancing techniques as a final stage in the MRIO database construction, any difference in row and column sums is removed from the associated value added figure. Thus, the original data is changed as little as possible. The authors use matrix difference statistics to identify the variation between the RASed and non RASed versions of the database. Findings suggest high correlation between the balanced and unbalanced versions of the economic matrices and lower when emissions results matrices are calculated.

\(^{11}\)Purchasing Price Parity adjusts the prices of goods and services to represent the same volume of goods regardless of the country of purchase. It allows the relative value of currencies to be determined.
2.5.1.3 Monte Carlo Techniques

Monte Carlo methods involve propagating repeated random input variables through a calculation and observing the effect on the output. They have proved to be useful in estimating the SD of MRIO multipliers and work by the generation of thousands of versions of the MRIO table being created which contain random, normally or log-normally distributed adjustments to the cells of the original matrix. A matrix representing the difference between the original matrix and each of the randomly generated adjustments \((\text{MRIO} - \text{MRIO}')\) has a mean of zero and the total relative SD of the combined input variables. Each of the thousands of newly generated tables is then subjected to the matrix calculation and the change in multipliers can be observed. Recently, Monte Carlo techniques have been used to estimate an 89% probability that the UK’s carbon footprint increased between 1994 and 2004 (Lenzen et al. 2010) and to show that while uncertainties around the total Dutch carbon footprint are low, lower tiered impacts attributed at the regional and sector level contained higher uncertainty (Wilting 2012).

Moran and Wood (2014) use Monte Carlo methods to perturb each cell of the emissions vector; interactions matrix and matrix of final demand in each of Eora, EXIOPOL, GTAP and WIOD by up to 10% to investigate whether there is convergence in the CBA of the databases. The authors also repeat the process using the same emissions databases with each model. This is described as harmonising the satellite account. The study assesses whether the range of CBA outcomes for each country for each model overlap the multi-model mean. Moran and Wood (2014) find that even after harmonising the emissions vector for many countries, the difference between model results is larger than one standard deviation.

2.5.2 Calculated Differences in CBA of Eora, GTAP and WIOD

The techniques described in Sect. 2.5.1 concentrate on taking a single MRIO database and quantifying the effect of a change in either the source data or construction on the resulting CBA. None of the techniques described above quantify how differences between the CBA calculated by different databases can be related to the differences in their construction. This book exploits this research gap by identifying techniques to understand difference and attempt to trace difference back to the MRIO source data and construction metadata as described in Sects. 2.3.1–2.3.3.

Table 2.3 shows the CBA in MtCO\(_2\) as calculated by Eora, GTAP and WIOD for the year 2007. The CBA calculated here includes the emissions associated with a country’s demand for products and the direct domestic household emissions from home heating and private transportation. Each account is compared to the mean account and the percentage difference is shown. There is clearly considerable variation in the outcomes with Luxembourg in particular having a very wide variation.
### Table 2.3
Consumption-based accounts (CBA) for 2007 in MtCO\(_2\) as calculated by Eora, GTAP and WIOD and deviation from the mean. Here the CBA includes direct emissions from households

<table>
<thead>
<tr>
<th>Country</th>
<th>Eora (%</th>
<th>GTAP (%)</th>
<th>WIOD (%)</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>434</td>
<td>4.9</td>
<td>347</td>
<td>−15.7</td>
</tr>
<tr>
<td>Austria</td>
<td>105</td>
<td>5.2</td>
<td>92</td>
<td>−7.8</td>
</tr>
<tr>
<td>Belgium</td>
<td>116</td>
<td>−17.1</td>
<td>157</td>
<td>11.9</td>
</tr>
<tr>
<td>Bulgaria</td>
<td>44</td>
<td>5.2</td>
<td>40</td>
<td>−4.1</td>
</tr>
<tr>
<td>Brazil</td>
<td>425</td>
<td>12.9</td>
<td>338</td>
<td>−10.1</td>
</tr>
<tr>
<td>Canada</td>
<td>543</td>
<td>−1.5</td>
<td>531</td>
<td>−3.6</td>
</tr>
<tr>
<td>China</td>
<td>4,840</td>
<td>6.9</td>
<td>4,174</td>
<td>−7.8</td>
</tr>
<tr>
<td>Cyprus</td>
<td>14</td>
<td>7.4</td>
<td>14</td>
<td>4.4</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>114</td>
<td>7.4</td>
<td>93</td>
<td>−12.1</td>
</tr>
<tr>
<td>Germany</td>
<td>948</td>
<td>−2.0</td>
<td>896</td>
<td>−7.4</td>
</tr>
<tr>
<td>Denmark</td>
<td>77</td>
<td>−3.6</td>
<td>84</td>
<td>6.5</td>
</tr>
<tr>
<td>Spain</td>
<td>472</td>
<td>3.6</td>
<td>415</td>
<td>−8.9</td>
</tr>
<tr>
<td>Estonia</td>
<td>21</td>
<td>7.9</td>
<td>19</td>
<td>−2.1</td>
</tr>
<tr>
<td>Finland</td>
<td>81</td>
<td>3.1</td>
<td>74</td>
<td>−5.4</td>
</tr>
<tr>
<td>France</td>
<td>610</td>
<td>5.2</td>
<td>542</td>
<td>−6.6</td>
</tr>
<tr>
<td>Great Britain</td>
<td>830</td>
<td>5.0</td>
<td>751</td>
<td>−4.9</td>
</tr>
<tr>
<td>Greece</td>
<td>162</td>
<td>0.6</td>
<td>168</td>
<td>4.4</td>
</tr>
<tr>
<td>Hungary</td>
<td>70</td>
<td>4.1</td>
<td>60</td>
<td>−10.4</td>
</tr>
<tr>
<td>Indonesia</td>
<td>352</td>
<td>1.4</td>
<td>336</td>
<td>−3.3</td>
</tr>
<tr>
<td>India</td>
<td>1,286</td>
<td>−1.3</td>
<td>1,252</td>
<td>−3.9</td>
</tr>
<tr>
<td>Ireland</td>
<td>61</td>
<td>−4.9</td>
<td>59</td>
<td>−7.4</td>
</tr>
<tr>
<td>Italy</td>
<td>611</td>
<td>2.9</td>
<td>549</td>
<td>−7.4</td>
</tr>
<tr>
<td>Japan</td>
<td>1,482</td>
<td>8.9</td>
<td>1,232</td>
<td>−10.3</td>
</tr>
<tr>
<td>Korea</td>
<td>595</td>
<td>10.1</td>
<td>474</td>
<td>−12.2</td>
</tr>
<tr>
<td>Lithuania</td>
<td>27</td>
<td>9.4</td>
<td>19</td>
<td>−20.9</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>19</td>
<td>24.6</td>
<td>17</td>
<td>12.5</td>
</tr>
<tr>
<td>Latvia</td>
<td>14</td>
<td>−6.1</td>
<td>17</td>
<td>7.5</td>
</tr>
<tr>
<td>Mexico</td>
<td>450</td>
<td>3.1</td>
<td>416</td>
<td>−10.1</td>
</tr>
<tr>
<td>Malta</td>
<td>5</td>
<td>10.3</td>
<td>4</td>
<td>−1.2</td>
</tr>
<tr>
<td>Netherlands</td>
<td>184</td>
<td>0.9</td>
<td>191</td>
<td>−6.8</td>
</tr>
<tr>
<td>Poland</td>
<td>309</td>
<td>9.5</td>
<td>282</td>
<td>−9.5</td>
</tr>
<tr>
<td>Portugal</td>
<td>83</td>
<td>8.7</td>
<td>72</td>
<td>−8.7</td>
</tr>
<tr>
<td>Romania</td>
<td>108</td>
<td>6.4</td>
<td>91</td>
<td>−16.3</td>
</tr>
<tr>
<td>Russia</td>
<td>1,246</td>
<td>5.6</td>
<td>1,236</td>
<td>−4.9</td>
</tr>
<tr>
<td>Slovakia</td>
<td>60</td>
<td>37.1</td>
<td>37</td>
<td>−19.6</td>
</tr>
<tr>
<td>Slovenia</td>
<td>20</td>
<td>4.8</td>
<td>19</td>
<td>−9.1</td>
</tr>
<tr>
<td>Sweden</td>
<td>94</td>
<td>4.8</td>
<td>82</td>
<td>−10.2</td>
</tr>
<tr>
<td>Turkey</td>
<td>321</td>
<td>6.0</td>
<td>306</td>
<td>−10.5</td>
</tr>
<tr>
<td>Taiwan</td>
<td>162</td>
<td>−15.1</td>
<td>189</td>
<td>−3.7</td>
</tr>
<tr>
<td>United States</td>
<td>6,662</td>
<td>8.5</td>
<td>6,089</td>
<td>−7.2</td>
</tr>
<tr>
<td>All industry emissions</td>
<td>28,237</td>
<td>11.1</td>
<td>22,800</td>
<td>−10.3</td>
</tr>
<tr>
<td>All direct household emissions</td>
<td>2,194</td>
<td>−33.3</td>
<td>3,724</td>
<td>13.1</td>
</tr>
<tr>
<td>TOTAL</td>
<td>30,431</td>
<td>6.0</td>
<td>26,524</td>
<td>−7.6</td>
</tr>
</tbody>
</table>
in estimates. This finding is also identified by Moran and Wood (2014). Eora tends to give estimates of CO$_2$ CBA that are larger than the mean and GTAP underestimates when compared to the mean. There is also considerable difference in the emissions designated to industries and those for households with Eora’s household estimate nearly 2,000 MtCO$_2$ lower than that of GTAP and Eora. As described in Sect. 2.3, Eora takes the territorial principle to emissions allocation. The emissions for industries therefore show greater difference than the global total difference. The techniques used in this book will focus mainly on the differences in the MRIO databases, meaning that the industrial emissions are of particular interest.

Figure 2.9 displays the differences in CBA graphically. The CBA is converted to tonnes CO$_2$ per capita figures for ease of display. In Fig. 2.9 the values are split by direct household emissions and emissions allocated to products. Direct household emissions are shown by the darker parts of each bar. Figure 2.9 clearly shows that for each country Eora has a lower estimate of direct household emissions.

While results were being compiled for this book, and also the Owen et al. (2014) submission to the ESR special edition on MRIO comparisons, Arto et al. (2014) independently produced a study comparing GTAP and WIOD. Their research compares the data sources used by both databases and gives some detail of the construction technique. A weighted relative percentage difference is calculated for a common classification versions of the GTAP and WIOD intermediate interaction matrices, final demand matrices and emissions vectors to assess the similarity between the building blocks of each database. Arto et al. (2014) also use decomposition methods (see Sect. 2.8), as this study does (see Owen et al. 2014), to attribute the difference in CBA as calculated by GTAP and WIOD to the final demand vector, interactions matrix, emissions vector and total output vector. The findings of this similar study will be addressed in Chap. 5 of this book, but it should be noted that this study compares just two databases and there is little attempt to relate the differences back to the accompanying metadata or to comment on how these differences might affect the use of model outcomes in policy.

As Dietzenbacher et al. (2013, p. 73) state, “one database should not be seen as ‘better’ than another database” and it will not be the intention of this book to declare one database the most accurate. Rather, the intention is to explore techniques to help identify and quantify the differences and the reason for the differences shown in Table 2.3 and Fig. 2.9. Where Dietzenbacher et al. (2013, p. 73) embrace the difference in MRIO databases and their construction because one might “be better (or more appropriate) for answering some questions but not for other questions”, Moran and Wood (2014, p. 246) suggest that with “continued improvements in modelling [the databases will converge] towards the underlying correct statistical account and that convergence of results is better than divergence”. Such viewpoints will be explored in the discussion and conclusion sections of this book.
2.5 Differences in MRIO Outcomes

Fig. 2.9 Differences in per capita CO₂ CBA for the 40 common countries in Eora, GTAP and WIOD. Bars split by embodied product emissions (lighter) and direct household emissions (darker).
2.6 Policy Applications, Level of Detail and Uncertainty

The results from MRIO databases can be used at a variety of scales from national level CBA, to sector level footprints down to identifying the contribution of a particular sector, from a particular country in a good’s production chain (Peters 2010). The confidence associated with results generally reduces as the scale gets finer and more detailed. This is because, as described above, the creation of the off-diagonal trade portions of MRIO tables requires some level of estimation meaning that values at the cell-by-cell level are uncertain. Rather than review the use of MRIO outcomes for all policy applications, this section of the literature review approaches the question from the concept of scale and comments on the reliability of evidence that could be potentially be used for policy.

2.6.1 National CBAs

The calculation of a national CBA requires the sum of a national level results matrix and it has been shown that regardless of sector and region aggregation, national level footprint remain fairly stable (Andrew et al. 2009; Peters and Solli 2010) thus, this calculation is the most robust of those discussed in this section. There are numerous examples of MRIOs being used for CBA measures including: the carbon footprint of nations (Hertwich and Peters 2009) and the water footprint of nations (Feng et al. 2011), both calculated using GTAP; and the material footprint of nations (Wiedmann et al. 2015) and the employment footprint of nations (Alsamawi et al. 2014), both calculated using Eora. See Wiedmann (2016) for an overview. Barrett et al. (2013) and Wiedmann and Barrett (2013), use the UK as a case study and explain the role national CBAs could have in policy by being an alternative indicator to be reported alongside territorial emissions. Barrett et al. (2013) demonstrate that Eora, GTAP and the UKMRIO report different CBA for the period 1990–2009 but the underlying trend in the consumption-based CO₂ emissions trajectory is similar. Before adopting the CBA as an indicator, the UK government requested an investigation into the robustness of the results, which led to the Monte Carlo analysis described previously (Barrett et al. 2013; Lenzen et al. 2010).

If CBA are reported over a time series, investigation of the year-on-year drivers of change can be a useful policy application. For example, Baiocchi and Minx (2010) demonstrate that the UK government’s Sustainable Development Strategy, which aims to improve the emissions efficiency of production, may not be enough to curb emissions in the face of increasing rises in the demand for goods. To decompose CBA results into drivers usually requires the exclusion of the effect of prices. WIOD is the only MRIO database thus far to report tables in previous years prices allowing the price effect to be eliminated (see Sect. 2.8 for further discussion of decomposition). Brizga et al. (2014) use WIOD to show that final demand is the dominant driver of the increase in the emissions CBA in three Baltic states from 1995–2009.
2.6 Policy Applications, Level of Detail and Uncertainty

2.6.2 Identifying the Imported Component of CBA

Splitting the CBA into those emissions where the source is domestic and those which are imported from abroad requires a further level of detail. Understanding the role of trade in global emissions has great policy relevance when considering producer versus consumer responsibility in GHG emissions reduction targets (Lenzen et al. 2007). However, Barrett et al. (2013) warn that CBA are not the solution to climate policy and should be seen as providing complementary and alternate information to the producer/territorial account.

Davis and Caldeira (2010) were the first to assign a figure to the proportion of global CO₂ emissions that were traded. Using the GTAP MRIO from 2004, they find that in wealthy nations more than 30% of the CBA is made up of imported emissions. Peters et al. (2011a) also use the GTAP MRIO to calculate the portion of Global CO₂ emissions that were associated with trade and to show that this portion grew between 1990 and 2008. However, the authors include considerable discussion of the uncertainties inherent in their calculation in the supporting information accompanying the manuscript. Since GTAP is not available as a continuous time series (see Table 2.2), data from 1997 was used as the trade balance for the time period 1990–1999, 2001 for 2000–2002 and 2004 for 2003–2008. Finding the sum of domestic and imported emissions requires summing across the rows of the national results table. This calculation should be fairly robust since, looking back to the construction methods explained in Sects. 2.2–2.3, the domestic and imports split is a fundamental element of the base building block—the SRIO table.

Many of the ‘footprint of nations’ studies have also commented on the role of trade. For example, using Eora, Wiedmann et al. (2015), when investigating the material footprint of nations, find that the material impact of imported goods is around three times the size of the physical quantity of the good itself. Similarly, Simas et al. (2014) use EXIOBASE to determine the labour impacts embedded in trade.

2.6.3 Impact by Source Nation and/or Product Destination

A further level of detail is to break down a nation’s CBA either to show the source nation and industry of the emissions or to show the final product footprint. Wiedmann et al. (2011) explain that product footprints may become policy relevant if eco-labelling becomes a requirement of product sustainability standards. Breaking the CBA down to show source nation and industry requires summing the relevant rows of a national results table. The BTD was used to break down imports by industry and country so this summation should be reasonably accurate. On the other hand, product footprints require column sums. As Sects. 2.2 and 2.3 explain, BTD is disaggregated across the off-diagonal matrices because the destination (or rather end product) is not recorded in the BTD statistics. This means that product footprints should be treated with less certainty than source footprints.
As early as 2010, Davis and Caldeira (2010) reported the breakdown of CBAs by product using GTAP 2004. More recently, Alsamawi et al. (2014) have analysed the employment footprint in traded goods and shown ranked lists of each countries’ imports by commodity and place of origin. The authors propose that developing countries have a large workforce involved in the production of electronics, agriculture and chemicals that that support the lifestyles of richer nations.

2.6.4 Supply Chain Analysis

Finally, the identification of an individual cell in a region’s CBA result table can reveal for each product, the proportion of product footprint that is sourced from each sector by import region. This level of detail has high uncertainty attached to it since the value is generated as the product of a number of assumptions. Nevertheless, Lenzen et al. (2012b) when analysing the land use impact associated with imported goods to understand the biodiversity impacts of trade, use the proportion of the land footprint of German coffee that is from Mexican agriculture to estimate how Germany’s coffee consumption can be linked to the threatened the habitat of the Mexican spider monkey.

As explained in Sect. 2.4.2, the OECD is in the process of developing a more comprehensive bilateral trade database which may improve the accuracy of the off-diagonal matrices. This means that the OECD ICIO can start to instigate projects investigating global value chains, such as Rouzet and Miroudot (2013) and the OECD-WTOs TiVA (Trade in Value Added) initiative. TiVA aims to calculate “the value added by each country in the production of goods and services that are consumed worldwide” (OECD 2014b).

It is clear that there is considerable work to do in assessing the difference between MRIO databases; identifying the cause of difference and commenting on how this uncertainty might have implications for the use of MRIO outcomes in policy. Sections 2.7–2.10 of the literature review are dedicated to reviewing techniques that can be used to understand difference.

2.7 Matrix Difference Statistics

Matrix difference statistics can be used to measure how different two matrices are from each other. Knudsen and Fotheringham (1986) identify three types of matrix difference statistics: distance statistics; goodness-of-fit; and information-based statistics. Distance statistics measure the cell-by-cell deviations between the two matrices and then calculate a single value as a description of the overall difference. Goodness-of-fit calculations measure how well the two matrices correlate to each other. And finally, information-based statistics compare the probability distributions of the result matrices. Information theory is concerned with the quantification of information.
2.7 Matrix Difference Statistics

Knudsen and Fotheringham (1986). Each type of statistic measures a different facet of how two matrices could be described as being similar to each other, therefore to gain a full understanding of how close two matrices are, several statistical measures should be used. In fact, Butterfield and Mules (1980, p. 293) state that “there exists no single statistical test for assessing the accuracy with which a matrix corresponds to another” and there are numerous examples in the literature of authors using, a suite of matrix comparison statistics in their work (Gallego and Lenzen 2006; Günlük-Senesen and Bates 1988; Harrigan et al. 1980; Knudsen and Fotheringham 1986). More detail on the specific matrix difference statistics chosen for this study is given in Sect. 3.2 along with justification for their selection.

In the years before readily available IO tables, analysts often estimated data tables for year $t_1$ based on year $t_0$ tables. With limited new data available, for certain elements of the table, RAS balancing techniques were applied to update missing values and ensure a balanced table. Once the tables for $t_1$ had been released, analysts could use matrix difference statistics to explore the accuracy of the observed and estimated tables (McMenamin and Haring 1974). Similarly, analysts have estimated sub-regional IO tables from national tables and then used difference statistics to examine the reliability of their estimates (Harrigan et al. 1980; Jackson and Comer 1993; Morrison and Smith 1974). Finally, matrix difference statistics have been used to measure the variation between pre and post RAS transaction matrices to further understand the effect of balancing techniques (Gallego and Lenzen 2006; Geschke et al. 2014; Wiebe and Lenzen 2016). Beyond the field of input–output analysis, Knudsen and Fotheringham (1986) employ comparison statistics when investigating a model that predicts flows. The actual and predicted flow matrices are compared and the difference evaluated using a number of comparison statistics.

As described above, there are many examples of matrix difference statistics being used with IO databases. The statistics are used to compare estimated and actual tables and to look at the effect of construction techniques, such as RAS balancing. These examples exclusively consider the difference between two tables from the same database. There are no examples of matrix difference statistics being used to understand the variation between different MRIO databases—a gap in this field of research.

2.8 Structural Decomposition Analysis

Decomposition analyses are used to understand changes in economic, environmental and other socio-economic indicators over time (Hoekstra and van der Bergh 2003). To decompose change at the sector level, two techniques are commonly employed: structural decomposition analysis (SDA) and index decomposition analysis (IDA). Hoekstra and van der Bergh (2003) explain that SDA uses the IO framework, whereas IDA calculates change using aggregated sector information. This means that SDA is able to identify the effects of a change in the technical requirements matrix and also to understand the effects of alterations in final demand—both of
which are not possible using IDA techniques. This study will use SDA techniques
to determine the difference between MRIO databases because the differences due to
demand and the technical requirement matrix may be significant in this type analysis.
Thus, the remainder of this section draws mainly from the SDA literature.

Structural decomposition analysis (SDA) is an “analysis of economic change
by means of a set of comparative static changes in key parameters in an input–
output table” (Rose and Chen 1991, p. 3). SDA takes the component parts of the
fundamental Leontief equation and calculates the effect each term (or determinant)
has on the change in consumption-based account For example, an SDA can isolate
and estimate the effect of technological change, the technology mix and level of
demand on a year-on-year change in a CBA (Rose and Casler 1996). In some cases,
when the total effects of all the determinants do not equal the total observed change,
a residual has to be calculated. There are two types of decomposition calculations:
additive and multiplicative (Rose and Casler 1996). The additive type decomposes the
difference between time $t$ and time $t + 1$ into several determinant effects, whereas the
multiplicative type decomposes the relative growth into determinant effects Hoekstra
and van der Bergh (2003). Hoekstra and van der Bergh (2003, p. 43) state that “the
reason to choose the additive or multiplicative decomposition is generally a matter
of presentation” and that “non-experts interpret additive decompositions relatively
easily”. This book chooses to explore additive SDA for two reasons: firstly, because
of its ease of interpretation and secondly because the concept of ‘growth’ makes
little sense when comparing two databases. The following text therefore concentrates
exclusively on additive SDA techniques and applications.

There are several different methods that can be used to calculate additive SDA.
One of the main reasons that there are so many techniques is that the calculation
assigns indexes (or weights) to each determinant and there is no single way of deter-
mining what those weights should be (Hoekstra and van der Bergh (2003)). Ang
(2004) distinguishes two methods for assigning indices: by percentage change and
by logarithmic change. Methods of assigning weight to determinants that are based
on Laspeyres decomposition use percentage change; whereas other Divisia rooted
techniques use logarithmic changes. Again, ease of interpretation is one of the rea-
sons why analysts prefer one technique over another and the percentage change is
easier to understand (Ang 2004). However, Divisia rooted methods are described by
Ang (2004, p. 1133) as “being more scientific”. This is because if a change of 20 to 40
is observed between times $t_0$ and $t_1$, this can either be described as a 100% increase
from $t_0$ to $t_1$ or a 50% decrease from $t_1$ to $t_0$ (Ang 2004). A log percent change records
the changes in both directions as 69.3% but this is more complicated to relate back to
the original numbers.\textsuperscript{12} When deciding which of the additive SDA techniques to use
in this study, Hoekstra and van der Bergh’s (2003) classification of the properties of
indices is useful. The authors describe three properties of a decomposition technique:

\begin{align*}
\ln \frac{20}{10} &= 0.693 \\
\ln \frac{10}{20} &= -0.693.
\end{align*}
Table 2.4  Features of the main additive SDA techniques (adapted from Hoekstra and van der Bergh 2003)

<table>
<thead>
<tr>
<th>Technique</th>
<th>Percent weights or logarithmic weights</th>
<th>Completeness</th>
<th>Time reversal</th>
<th>Zero value robustness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laspeyres</td>
<td>Percent</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Marshall-Edgeworth</td>
<td>Percent</td>
<td>Only in 2 determinant case</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Paasche</td>
<td>Percent</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Conventional Divisia</td>
<td>Percent</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Log-Mean Divisia</td>
<td>Logarithmic</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes if small number replaces zeros</td>
</tr>
<tr>
<td>Adaptive Weighting Divisia</td>
<td>Logarithmic</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Shapely-Sun</td>
<td>Percent</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Dietzenbacher and Los</td>
<td>Percent</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

- Completeness—the decomposition has a residual of zero
- Time reversal—if the order is reversed is the same result calculated
- Zero value robustness—if logarithms are involved in the calculations, this causes an issue when there are zeros in the dataset

For comparison of two different MRIO tables rather than the same MRIO for two years, the time reversal property becomes very important. The same result should be calculated when comparing GTAP to WIOD as found comparing WIOD to GTAP. Table 2.4 compares additive SDA techniques.

Hoekstra and van der Bergh (2003) explain that the Laspeyres, Marshall-Edgeworth, Paasche, Conventional Divisa and Adaptive Weighting Divisia decomposition techniques fail on at least one of these properties. This leaves the Log-Mean Divisia Index\(^{13}\) (LMDI) (Ang and Choi 1997), the Shapely-Sun\(^{14}\) (S-S) (Sun 1998) and the Dietzenbacher and Los (D&L) (Dietzenbacher and Los 1998) techniques. In the following section we shall explore each of these approaches.

\(^{13}\)Known as the ‘Refined Divisia’ technique in Ang and Choi (1997) and Hoekstra and van der Bergh (2003).

\(^{14}\)Known as the ‘Sun’ technique in Hoekstra and van der Bergh (2003).
2.8.1 Log-Mean Divisia Index

LMDI tends to be used for IDA rather than SDA and was first proposed by Ang and Choi (1997) as a ‘refined divisia’ method. Where other techniques use arithmetic mean weights and require a residual in the calculations, the LMDI method uses a logarithmic mean weight and decomposes perfectly. The authors also show that if any zeros in the dataset are replaced by near zero values, the decomposition converges to a result. Ang (2004) goes as far as to recommend that the LMDI technique is the most appropriate decomposition method for policy making in energy.

2.8.2 Shapely-Sun

Sun (1998) proposed a refined Laspreyes decomposition technique that removed the need for a residual term. In Laspreyes decompositions, the residual term can be described to be the effect of the interaction of a number of determinants. Sun (1998) demonstrates how this interaction effect can be reassigned and equally split among the main residual effects (Ang 2004; Hoekstra and van der Bergh 2003). The Sun (1998) technique was shown to be identical to a method proposed by Shapley\(^{15}\) and so this method is now referred to as the Shapley-Sun (S-S) technique (Ang et al. 2003; Ang 2004).

2.8.3 Dietzenbacher and Los

The D&L decomposition technique does not calculate a single index but rather develops a range of indices with no residual term (Dietzenbacher and Los 1998; Hoekstra and van der Bergh 2003). For example, if the environmentally extended Leontief equation is the product of three terms there are a total of six, \(3! = 6\), decomposition equations that can be formulated to describe the change in CBA (see Sect. 3.3.1 for further details). This means that there is no unique solution and each of the decomposition forms is equally valid (Dietzenbacher and Los 1998). The mean of each of the decomposition solutions is often taken as an indication of the influence of each determinant but Dietzenbacher and Los (1998) note that the maximum, minimum and standard deviation of each determinant can and should be reported.

Hoekstra and van der Bergh (2003) suggest that the mean effect of all of the D&L indices is the same result as the indices calculated for S-S decomposition. This is later proved by de Boer (2009).

\(^{15}\text{For details of the Shapley method see Albrecht et al. (2002).}\)
2.8.4 Applications of Structural Decomposition Analysis

The use of additive SDA to understand the drivers of emissions change over time is well documented. Studies investigating the causes of a nation’s increase in carbon CBAs include Baiocchi and Minx (2010), Guan et al. (2008, 2009), Minx et al. (2011), Peters et al. (2007), Tian et al. (2014). Interestingly each of these studies employs additive D&L methods. Both Baiocchi and Minx (2010) and Minx et al. (2011) report the calculated ranges in the effect of each determinant as suggested by Dietzenbacher and Los (1998). However, comment on the minimum, maximum, and or variance of the effect of each term is not commonly found in the SDA literature. LMDI techniques seem to be more popular in studies decomposing changes in energy (see for example Wachsmann et al. 2009).

Ang and Zhang (2000) show that there are very few examples of SDA being used for anything other than an assessment of the drivers of change over time; their survey of 124 decomposition studies find just three that do this. Jakob and Marschinski (2013), however, demonstrate how the S-S technique can be used to understand trade balances. Rather than finding the difference in emissions between \(t_0\) and \(t_1\), the authors decompose the difference between a country’s exports and imports.

Dietzenbacher and Los (2000) warn that analyses that decompose a term such as total value added need to be treated with care due to the dependency problem. A decomposition equation containing three terms assumes each are independent of each other. The authors point out that “changes in intermediate input coefficient and in value added coefficient affect each other” (Dietzenbacher and Los 2000, p. 4). SDA applied to measures of consumption-based emissions require the calculation of the emissions per unit of output and this dependency issue will need to be considered. 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 SDA 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.

This study is concerned with understanding the difference between the carbon CBAs as calculated by different MRIO databases. SDA provides a useful technique for considering the effect that each component of the environmentally-extended Leontief equation has on the difference in CBA. It is clear that there is a gap in the research for SDA to be used for this type of investigation. An understanding of the certainty of the effect of each component could prove very interesting. For example, if the effect of the difference in GTAP and WIOD’s final demand vectors is large but the variance in the size of this effect, as calculated by the D&L technique, is small, then there is a greater certainty that the difference in the CBA could be due to the final demand vector. If the variance is large, then the certainty of the importance of the effect is lessened. This study will therefore use the D&L method to calculate decompositions of CBA. Further details of the SDA equations themselves can be found in Sect. 3.3.
2.9 Structural Path Analysis

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 zeroth 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 steel production used to make a steel final demand product. A first order path has one further step in the supply chain: for example the emissions from steel production that are used to make cars for final demand. Most SPA studies rank these production chains or paths 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 (2006) state that there are very few IO studies that apply SPA and that hybrid life cycle assessment (LCA) techniques have been a more popular method employed to consider production chains. By 2015, this is still the case—SPA methods remain relatively unpopular. Wood and Lenzen (2003) use SPA and a 1995 SRIO database for Australia to compare the land use CBA of two Australian research institutions. Their analysis reveals a large proportion of the two institutions’ land use 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, Nitrous Oxide (NOX) and Sulphur Dioxide (SO$_2$) emissions as environmental extensions. Lenzen (2003) demonstrates that when considering energy and emissions rather than land use, the zeroth 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 having significant impact. This means that product chains will often have to be at least a first order chain to link to the land using sectors. There is significant direct energy and emissions use for a wider proportion of industrial sectors meaning that many zeroth order paths will be significant. The advantage of an emissions-based study is that the largest paths will be relatively short and quick to find during the SPA procedure. Both Lenzen (2003) and Peters and Hertwich’s (2006) analyses of Australia and Norway, respectively find that zeroth order paths involving electricity, metals, chemicals and transport services are significant.

Rather than look at all the production chains making up the entire emissions CBA, Acquaye et al. (2011) consider the upstream paths that contribute to the production of biofuels using a UK focused two-region MRIO database. The authors discuss how

---

16This is more commonly known as the ‘Ecological Footprint’ (see Wackernagel and Rees 1996).
SPA has been used in this case to identify carbon hot spots, or rather the highest carbon intensity path of the upstream supply chain or biodiesel.

It is clear that SPA is an underused technique in MRIO database analyses and as yet, there have been no SPA research published using Eora, GTAP and WIOD. In Sect. 2.10, a technique that uses SPA to compare year-on-year differences is discussed since it is the difference between the databases that is covered by this book.

2.10 Structural Path Decomposition

Structural Path Decomposition (SPD) was developed by Wood and Lenzen (2009) as a combination of SDA and SPA. Wood and Lenzen (2009) use SPD to understand changes in a production chain between two points in time. Where SDA assigns 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 production chain between \( t_0 \) and \( t_1 \) could occur in a zeroth order path such as the emissions involved in air travel as a result of the purchase of a plane ticket 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 can identifies which part of the chain has the highest difference associated with it. For example, in the first 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 LMDI calculated SDA technique for the SPD methodology and apply it to Australian SRIO tables for 1995 and 2005. There are no examples of other SDA methods—such as the D&L or S-S technique—used for SPD. The authors find that between 1995 and 2005, the largest changes in emissions production paths involved livestock and electricity. The element of the paths, which Wood and Lenzen (2009) name ‘the differential’ 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 CO\(_2\) emissions in Japanese supply chains between 1990 and 2000 and Gui et al. (2014) consider changes in CO\(_2\) 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 SDA 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 book may present the first application of SPD for this use. In addition there is also an
option to explore using the D&L or S-S SDA technique within the SPD calculations, which is considered more accurate than polar decompositions (de Boer 2009).

The equations used for SPA and SPD are presented in Sects. 3.4 and 3.5 respectively.

References


References


WIOD 2014. World Input Output Database.

Techniques for Evaluating the Differences in Multiregional Input-Output Databases
A Comparative Evaluation of CO2 Consumption-Based Accounts Calculated Using Eora, GTAP and WIOD
Owen, A.
2017, XXII, 217 p. 44 illus. in color., Hardcover
ISBN: 978-3-319-51555-7