Regional input-output models and the treatment of imports in the European System of Accounts (ESA)

Tobias Kronenberg

Springer

Regional input-output models and the treatment of imports in the European System of Accounts (ESA)

Tobias Kronenberg

Abstract Input-output models are useful tools for regional science owing to their ability to capture many of the distinguishing features of a regional economy. Since input-output tables are hard to find at the regional level, many researchers adopt non-survey techniques to derive regional input-output tables (RIOTs). Numerous methods have been suggested and this has spawned a stream of literature comparing the relative performance of these methods.

The present paper contributes to that literature by examining a largely neglected problem of nonsurvey techniques: the allocation of imports. In the European System of Accounts (ESA), there are two ways of allocating imports: inside the interindustry transactions matrix or outside. In the former case, they are allocated as imports in the sector that produces similar goods (indirect allocation). In the latter case, imported products are allocated to the sector that uses them (direct allocation).

This paper shows that the choice of a nonsurvey method should depend on the way in which imports are allocated. If the nonsurvey method is not properly chosen, the results of the procedure may be misleading and implausible. The paper concludes that LQ methods are better suited for regionalising input-output tables with directly allocated imports, whereas commodity-balance methods like CHARM are better suited for regionalising input-output tables with indirectly allocated imports.

Keywords Regional input-output model · Nonsurvey method · Location quotient · Commodity balance

JEL Classification C67 · R15

Dr. T. Kronenberg (✉)
Institut für Energieforschung, Systemforschung und Technologische Entwicklung,
Forschungszentrum Jülich GmbH, 52425 Jülich, Germany
e-mail: t.kronenberg@fz-juelich.de

Regionale Input-Output-Modelle und die Verbuchung der Importe im Europäischen System Volkswirtschaftlicher Gesamtrechnungen (ESVG)

Zusammenfassung Input-Output-Modelle sind für die Regionalwissenschaft nützlich, weil sie viele der Merkmale eines regionalen Wirtschaftsraums erfassen können. Da Input-Output-Tabellen auf regionaler Ebene schwer zu finden sind, werden häufig Nonsurvey-Verfahren angewendet, um regionale Input-Output-Tabellen (RIOT) zu erstellen. Zahlreiche Methoden sind vorgeschlagen worden. Angesichts der zunehmenden Methodenvielfalt ist eine Literatur entstanden, in der die Leistungen der unterschiedlichen Methoden miteinander verglichen werden.

Das vorliegende Papier trägt zu dieser Literatur bei, indem es sich einem weitgehend vernachlässigten Problem der Nonsurvey-Verfahren widmet: der Verbuchung der Importe. Im Europäischen System Volkswirtschaftlicher Gesamtrechnungen (ESVG) gibt es zwei Möglichkeiten, die Importe zu verbuchen: innerhalb der interindustriellen Transaktionsmatrix oder außerhalb. Im letzteren Fall werden importierte Produkte dem sie verwendenden Sektor zugeordnet (direkte Zuordnung). Im ersteren Fall werden sie als Importe dem Sektor, der ähnliche Güter produziert, zugeordnet (indirekte Zuordnung).

Dieser Artikel zeigt, dass die Wahl der Nonsurvey-Methode davon abhängen sollte, wie die Importe zugeordnet werden. Wenn die Nonsurvey-Methode nicht korrekt gewählt wird, kann dies zu irreführenden und implausiblen Ergebnissen führen. Der Artikel kommt zu dem Fazit, dass LQ-Methoden zur Regionalisierung von Tabellen mit direkt zugeordneten Importen geeignet sind, wohingegen Commodity-Balance-Methoden wie CHARM besser geeignet sind, um Tabellen mit indirekt zugeordneten Importen zu regionalisieren.

1 Introduction

Input-output analysis is widely used by authors working in the fields of regional science or regional economics. It is also becoming increasingly popular in environmental and ecological economics (Los 2011). Naturally, some ecological economists are also interested in conducting environmental impact studies for individual regions. Therefore, it is likely that regional input-output models will frequently be used for environmental impact studies in the future.

When studying a particular region, analysts often have to construct a regional input-output table (RIOT), since many statistical offices provide only national input-output tables (NIOTs). Fortunately, there are established methods for regionalising a NIOT and adapting it to regional characteristics (nonsurvey methods). Numerous methods have been suggested; the most recent ones are the FLQ (Flegg and Webber 1997, 2000; Flegg et al. 1995) and CHARM (Kronenberg 2009). A large and growing literature discusses the strengths and weaknesses of these methods (Bonfiglio and Chelli 2008; Morrison and Smith 1974; Richardson 1985; Schaffer and Chu 1969; Tohmo 2004).[1] However, the focus of the present paper is a different one.

---

[1]The earlier nonsurvey regionalisation methods were criticised for their neglect of cross-hauling (i.e. interindustry trade), which results in an underestimation of regional trade and an overestimation of re-

This paper aims at drawing attention to a crucial issue whose importance has not yet been fully realised in the literature on nonsurvey regionalisation methods: there are different variants of the symmetric input-output table (SIOT), and the choice of nonsurvey method should depend on the type of SIOT that is to be regionalised. The most important difference between the SIOT variants lies in the treatment of imported products. The United Nations handbook on input-output analysis identified four different variants, labelled alphabetically from "A" to "D" (United Nations 1973). This convention is also adopted in the present paper, and an additional variant "E" (for "Eurostat") is introduced to describe the tables based on the European System of Accounts (ESA 95).

A crucial finding of this paper is that location quotient (LQ) methods are suitable for variant B tables, whereas commodity balance (CB) methods are suitable for variant A and E tables. The existing literature has not paid much attention to this issue because regional economists mostly use variant B tables. Ecological economists, by contrast, are more likely to use variant A or E tables.

The paper proceeds as follows. Section 2 introduces definitions of variables and conventions on mathematical notation. Section 3 explains the different variants of the SIOT, largely following the exposition in the UN manual (United Nations 1973). Section 4 describes the interpretation of the coefficients derived from different tables. Section 5 identifies the implications that follow for those who want to construct a RIOT using nonsurvey methods. Finally, Sect. 6 provides some concluding remarks and suggests avenues for future research.
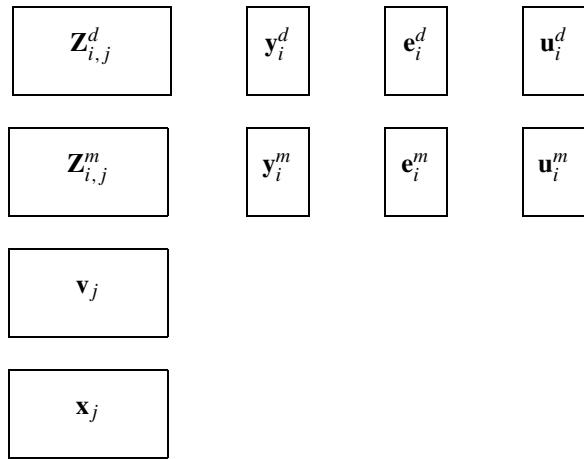
2 Definitions and conventions

Figure 1 shows the basic data which are needed to construct input-output tables of the sort that will be discussed below.

The following conventions will be used: the subscript i stands for products or commodities, whereas the subscript j stands for industries or homogeneous branches. As all the input-output tables provided by Eurostat are based on the concept of a "homogeneous branch", an artificial construct which by assumption produces only one type of commodity, the present paper also uses this concept. In reality, many firms produce more than one type of good, so it is impossible to assign the inputs and outputs of one firm unambiguously to one homogeneous branch. Therefore, the statistical offices "split" the firm and allocate its inputs and outputs over a number of homogeneous branches.[2] The details of this procedure are outside the scope of the present paper. In the following, the term "industry" will be used as a synonym for "homogeneous branch".

---

gional output multipliers (Boomsma and Oosterhaven 1992; Richardson 1985). For this reason, new approaches like the FLQ and CHARM were designed to allow for cross-hauling (Flegg and Tohmo 2011; Kronenberg 2009).

[2]Cf. http://circa.europa.eu/irc/dsis/nfaccount/info/data/esa95/en/een00101.htm: "Units of homogeneous production cannot usually be observed directly; data collected from the units used in statistical enquiries have to be re-arranged to form homogeneous branches".

Fig. 1 Basic data. Source: author's illustration

$$\boxed{\mathbf{Z}^d_{i,j}} \quad \boxed{\mathbf{y}^d_i} \quad \boxed{\mathbf{e}^d_i} \quad \boxed{\mathbf{u}^d_i}$$

$$\boxed{\mathbf{Z}^m_{i,j}} \quad \boxed{\mathbf{y}^m_i} \quad \boxed{\mathbf{e}^m_i} \quad \boxed{\mathbf{u}^m_i}$$

$$\boxed{\mathbf{v}_j}$$

$$\boxed{\mathbf{x}_j}$$

In Fig. 1 (and henceforth), a superscript d or m is used to indicate the origin of products (domestically produced or imported). Matrices are denoted by capital letters, vectors by lower case letters. Both are printed in bold type. The individual elements of a matrix are printed in italics. Thus, for example, $Z_{i,j}$ is element i, j of matrix $\mathbf{Z}$. It reports the total amount of product i used by industry j. The amount originating from domestic production is $Z^d_{i,j}$, whereas the amount imported is $Z^m_{i,j}$. Naturally, $Z^d_{i,j} + Z^m_{i,j} = Z_{i,j}$.

In addition to $\mathbf{Z}^d$ and $\mathbf{Z}^m$, the basic data table contains the following elements. The vector $\mathbf{v}$ reports value added (i.e. primary inputs) by industry, and the vector $\mathbf{x}$ reports output (i.e. production) by industry. The vectors $\mathbf{y}^d$ and $\mathbf{y}^m$ contain domestic final use of products, respectively. Domestic final use is defined as the sum of private consumption expenditure, public consumption expenditure, and gross capital formation. The vector $\mathbf{e}^d$ reports exports of domestically produced commodities, whereas $\mathbf{e}^m$ reports exports of imported commodities (i.e. re-exports). Finally, the vectors $\mathbf{u}^d$ and $\mathbf{u}^m$ describe the total use of domestically produced and imported commodities. Total use is defined as the sum of intermediate use, final domestic use, and exports. Mathematically:

$$u_i = \sum_{j=1}^{n} Z_{i,j} + y_i + e_i \tag{1}$$

where n is the number of products. Naturally, this relationship holds for both domestically produced and imported products:

$$u^d_i = \sum_{j}^{n} Z^d_{i,j} + y^d_i + e^d_i \tag{2a}$$

$$u^m_i = \sum_{j}^{n} Z^m_{i,j} + y^m_i + e^m_i \tag{2b}$$

Table 1 The comprehensive input-output table

	Homogeneous branches				Final uses		Total use
	1	...	n	Total	Domestic	Exports	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Domestically produced products							
1 1	$Z_{1,1}^d$	...	$Z_{1,n}^d$	r_1^d	y_1^d	e_1^d	u_1^d
2 ...	...	...	...	...	...	...	...
3 n	$Z_{n,1}^d$	...	$Z_{n,n}^d$	r_1^d	y_n^d	e_n^d	u_n^d
4 Subtotal	z_1^d	...	z_n^d	$z^d = r^d$	y^d	e^d	u^d
Imported products							
5 1	$Z_{1,1}^m$	...	$Z_{1,n}^m$	r_1^m	y_1^m	e_1^m	u_1^m
6 ...	...	...	...	...	...	...	...
7 n	$Z_{n,1}^m$	...	$Z_{n,n}^m$	r_1^m	y_n^m	e_n^m	u_n^m
8 Subtotal	z_1^m	...	z_n^m	$z^m = r^m$	y^m	e^m	u^m
9 Total interm. cons./final use	z_1	...	z_n	$z = r$	y	e	u
10 Value added	v_1	...	v_n	v			
11 Output	x_1	...	x_n	x			

Source: author's illustration

Figure 1 shows which data are needed for simple applications of input-output analysis. The big advantage of input-output tables is that they arrange these data in a straightforward manner that is consistent with standard book-keeping procedures. Table 1 shows how this can be done in the form of a comprehensive input-output table containing all of the relevant information.

The first column of Table 1 refers to the first industry (henceforth 'industry 1'). In the first three rows (with row 1 referring to product 1, row 3 to product n, and row 2 to "all products in between 1 and n"), the elements of $\mathbf{Z}^d$ concerning industry 1 are reported. Row 4 contains a subtotal, denoted by z_1^d. This is the sum of all domestically produced products that were used as intermediate inputs by industry 1. Below that, the relevant elements of $\mathbf{Z}^m$ are reported. Rows 5 to 7 show the use of imported products as intermediate inputs by industry 1, and row 8 contains the sum of these, z_1^m. In row 9, we find the sum of z_1^d and z_1^m. z_1 is the value of all intermediate inputs used by industry 1. Using these intermediate inputs, industry 1 generates a certain amount of value added, reported in row 10 and denoted by v_1. Taxes on products, which drive a wedge between basic prices and purchasers' prices and divert a share of value added to the government, are ignored here for the sake of simplicity. Depreciation (i.e. consumption of fixed capital) and taxes on production are also ignored. Under these assumptions, value added is simply the sum of compensation of employees (wages plus social security contributions, the reward for labour services) and net operating surplus (the remuneration of capital).

The following relationship holds by definition:

$$x_j = \sum_i^n Z_{i,j}^d + \sum_i^n Z_{i,j}^m + v_j \tag{3}$$

This expression states that the value of output produced by industry j is equal to the value of intermediate products used by that industry and the value added by that industry. This definition is in accordance with the classical theory of value, where firms buy inputs (commodities) of a certain value and generate additional value in the course of the production process. The added value is then distributed to the primary inputs labour, capital and land (although capital and land are unfortunately not displayed separately in the input-output tables).

Moving along the first row of Table 1, we can see how and where *domestically produced* products of type 1 are used. The first three columns show the amounts of product 1 that are used by industries 1 to *n* as intermediate inputs, and column 4 shows the sum of these. Column 5 shows the domestic final use (final consumption expenditure by households, NPISH,[3] and government as well as gross capital formation including stock formation) of product 1. In column 6 we observe the amount of product 1 that is exported to other countries. Finally, column 7 reports total use (i.e. the sum of intermediate use and final use). Rows 2 and 3 show the same for all other domestically produced commodities, and row 4 shows the sum of rows 1 to 3. Rows 5 through 8 show the same thing but for *imported* products. Thus, the comprehensive input-output table allows us to trace the use of domestically produced products (rows 1 through 4) separately from the use of imported products (rows 5 through 8). Row 9 is the sum of intermediate use (columns 1 to 4) and final use (columns 5 and 6). Row 10 reports value added by each industry, and row 11 reports the total output of each industry.

Eurostat does not supply comprehensive input-output tables of the type shown in Table 1. However, it does provide all the data that are required to produce such a table.[4] For almost every EU member country, there is an input-output table containing only domestically produced products, which contains data for $\mathbf{Z}_{i,j}^d$, $\mathbf{y}_i^d$, $\mathbf{e}_i^d$, $\mathbf{u}_i^d$, v_j, and $\mathbf{x}_j$. Furthermore, there is an 'import matrix', which contains data for $\mathbf{Z}_{i,j}^m$, $\mathbf{y}_i^m$, $\mathbf{e}_i^m$, and $\mathbf{u}_i^m$. Thus, with the available data we can produce a comprehensive input-output table. However, this is usually not done. Most input-output modellers prefer working with a symmetric input-output table (SIOT). What complicates matters is the fact that there are different ways of constructing a SIOT. These are discussed in the following section.

---

[3] NPISH = non-profit institutions serving households

[4] Input-output tables for EU member states and candidate countries are publicly available on the Eurostat website (http://epp.eurostat.ec.europa.eu/portal/page/portal/esa95_supply_use_input_tables/data/workbooks).

Table 2 Symmetric input-output table, variant 'A'

Products	Homogeneous branches				Final uses		Imports	Output
	1	...	n	Total	Domestic	Exports		
1	$Z_{1,1}$	...	$Z_{1,n}$	r_1	y_1	e_1	$-m_1$	x_1
...	...	...	...	...	...	...	...	...
n	$Z_{n,1}$	...	$Z_{n,n}$	r_n	y_n	e_n	$-m_n$	x_n
Total interm. use/final use	z_1	...	z_n	$z = r$	y	e	$-m$	x
Value added	v_1	...	v_n	v				
Output	x_1	...	x_n	x				

Source: author's illustration

3 Variants of the symmetric input-output table

Table 2 reports what we will call the SIOT Variant A.[5] At the core of the SIOT Variant A is an interindustry transactions matrix $\mathbf{Z}$, which reports the entire intermediate consumption of products (domestically produced and imported). Mathematically, $\mathbf{Z} = \mathbf{Z}^d + \mathbf{Z}^m$. Taking column sums of this matrix yields total intermediate consumption by industry, denoted by $\mathbf{z}$. Taking row sums yields total intermediate consumption by product, denoted by $\mathbf{r}$. By definition, summing $\mathbf{z}$ over j must yield the same result as summing $\mathbf{r}$ over i, so $z = r$ (but $\mathbf{z} = \mathbf{r}$ will usually not be true; it is possible but extremely unlikely).

In each homogeneous branch, intermediate consumption plus value added is equal to output:

$$\mathbf{x} = \mathbf{z} + \mathbf{v} \tag{4}$$

The bottom row of the SIOT variant A reports output by industry x_j. Total output is the sum of x_j over all $j : x = \sum_j^n x_j$.

Each row can be understood as a representation of the commodity balance. If a country uses more of product i than it produces, it must be a net importer of that product, and vice versa. In other words, net exports of product i must be equal to domestic output minus domestic use of that product. Mathematically:

$$e_i - m_i = x_i - r_i - y_i \tag{5}$$

Rearranging terms yields:

$$r_i + y_i + e_i - m_i = x_i \tag{6}$$

Going through row i of the SIOT variant A means going through (6). This is why imports are entered with a negative sign in the table.

---

[5]This is what Holub and Schnabl (1994) call "Variante A1". At the regional level, it is closely related to what Stäglin (2001) calls the "technological version".

Table 3 Symmetric input-output table, variant 'B'

Products	Homogeneous branches				Final uses		Output
	1	...	n	Total	Domestic	Exports	
1	$Z_{1,1}^d$	...	$Z_{1,n}^d$	r_1^d	y_1^d	e_1^d	x_1
...	...	...	...	...	...	...	...
n	$Z_{n,1}^d$	...	$Z_{n,n}^d$	r_1^d	y_n^d	e_n^d	x_n
Imported products	z_1^m	...	z_n^m	z^m	y^m	e^m	m
Total interm. use/final use	z_1	...	z_n	$z = r$	y	e	u
Value added	v_1	...	v_n	v			
Output	x_1	...	x_n	x			

Source: author's illustration

The symmetry of SIOT variant A is captured by the fact that each column represents the summation procedure resulting in the value of total output for each industry:

$$x_j = \sum_i^n Z_{i,j} + v_j \tag{7}$$

Thus, the final entry in row $i(x_i)$ is equal to final value in row $j(x_j)$ whenever $i = j$.

Table 3 shows the symmetric input-output table, variant B.[6] This variant is based on a different way of recording imports. In variant A, imports are allocated by product and the vector **m** consists of the m_i. In variant B, by contrast, imports are allocated by use, i.e. homogeneous branches and final users. z_j^m denotes the column sums of matrix $\mathbf{Z}^m$. Thus, z_j^m is the value of imported products that were used as intermediate inputs by industry j. Accordingly, z^m denotes the total value of imports used as intermediate inputs. y^m denotes the value of products that were consumed by final users in the country, and e^m denotes the value of imported products used for exports (re-exports). m denotes the total value of imports.

It is important to realise that $\mathbf{z}^m$ is very different from **m**. The latter is a column vector of length n (the number of different products), and element m_i is interpreted as "imported products of type i". The former is a row vector of length n, and element z_j^m is interpreted as "products of all types imported for use by industry j". Moreover, the sum of all elements is not equal—vector **m** contains all imported products, whereas vector $\mathbf{z}^m$ contains only those products imported for intermediate use, as imported products for final use are recorded elsewhere.

Table 4 shows how Eurostat currently compiles its symmetric input-output tables, including domestic products and imports according to the ESA 95 guidelines. This variant is called variant 'E' for 'Eurostat'. It is very similar to the type A table discussed above (Table 2); the only difference lies in the treatment of imports. At the core of Variant E is the interindustry transactions matrix **Z**, as in variant A.

---

[6]This is what Holub and Schnabl (1994) call "Variante B" and Stäglin (2001) calls "regional version".

Table 4 Symmetric input-output table, variant 'E'

Products	Homogeneous branches				Final uses			Total use
	1	...	n	Total	Domestic	Exports	Total	
1	$Z_{1,1}$	...	$Z_{1,n}$	r_1	y_1	e_1	f_1	u_1
...	...	...	...	...	...	...	...	...
n	$Z_{n,1}$	...	$Z_{n,n}$	r_n	y_n	e_n	f_n	u_n
Total interm. use/final use	z_1	...	z_n	$z = r$	y	e	f	u
Value added	v_1	...	v_n	v				
Output	x_1	...	x_n	x				
Imports of similar goods	m_1^E	...	m_n^E	m				
Total supply	s_1	...	s_n	s				

Source: author's illustration

For the type E table, an additional vector labelled "total supply" (denoted by **s**) is introduced. It is defined as the sum of domestic production and imports of similar commodities:

$$\mathbf{s} = \mathbf{x} + \mathbf{m}^E \tag{8}$$

The row vector $\mathbf{m}^E$ refers to imports by commodity. $\mathbf{m}^E$ stands for 'vector of imports constructed the Eurostat way'. m_j^E is the value of imports of commodity j, not the value of products imported by industry j. This is a crucial difference, as we will see below. Mathematically, $\mathbf{m}^E$ is the transpose of $\mathbf{m}$.

Total final use is defined as the sum of domestic final use and exports:

$$\mathbf{f} = \mathbf{d} + \mathbf{e} \tag{9}$$

Total use is equal to the sum of intermediate use (by product) and final use:

$$\mathbf{u} = \mathbf{r} + \mathbf{f} \tag{10}$$

Finally, it is true by definition that

$$\mathbf{s} = \mathbf{u} \tag{11}$$

Thus, the IOT is symmetric in the sense that $s_j = u_i$ when $i = j$.

To sum up, there is not a big difference between variants A and E. Actually, variant A can easily be converted into variant E by simply transposing the import vector and adjusting the column/row totals accordingly (and vice versa). There is, however, a great difference between variants E and B, because the treatment of imported products is very different. The implications of this will be discussed in the following section.

4 Interpretation of coefficients

When input-output tables are used in empirical research, it is common practise to compute input-output coefficients to describe the structure of the economy, to identify supply chains involving various industries, and to provide the basis for model-based impact studies.[7] Therefore, it is important to understand correctly the interpretation of the input-output coefficients that can be computed from the different tables.

In variant A, input-output coefficients are defined as:

$$a_{i,j}^A = \frac{Z_{i,j}}{x_j} \qquad (12)$$

These coefficients describe how many units of input i were used/needed to produce one unit of output j. Therefore, they can be interpreted as *technological coefficients*.

In variant B, by contrast, input-output coefficients are defined as:

$$a_{i,j}^B = \frac{Z_{i,j}^d}{x_j} \qquad (13)$$

These coefficients do not tell us how many units of input i were used to produce one unit of output j, because they refer only to those inputs that were produced domestically. Imported inputs are ignored. To make this point clearer, let us define a trading coefficient $t_{i,j}$ as:

$$t_{i,j} = \frac{Z_{i,j}^d}{Z_{i,j}} \qquad (14)$$

$t_{i,j}$ is the share of input i used by industry j that originates from domestic production. Conversely, $(1 - t_{i,j})$ can be interpreted as the import share. Using (12), (13) and (14), we can write the relationship between $a_{i,j}^A$ and $a_{i,j}^B$ as:

$$a_{i,j}^B = t_{i,j} a_{i,j}^A \qquad (15)$$

Thus, $a_{i,j}^B$ will generally be smaller than $a_{i,j}^A$. The $a_{i,j}^B$ coefficients differ from the true technological coefficients ($a_{i,j}^A$) owing to international trade. Therefore, they cannot be interpreted as technological coefficients. They are a mixture of technology and trade.

Variant E is interesting because it allows the calculation of two different kinds of coefficient. If we divide the elements of **Z** by output, we are performing the same calculation as in (12). Mathematically:

$$a_{i,j}^E = \frac{Z_{i,j}}{x_j} = a_{i,j}^A \qquad (16)$$

---

[7]Input-output coefficients serve as the basis for a variety of models, ranging from the very simple open static quantity model to more sophisticated models that allow for endogenous consumption, investment, flexible prices, and technological change.

Thus, variant E also makes possible the computation of technological coefficients. However, if we divide $Z_{i,j}$ by s_j, we get a different kind of coefficient (an input-*supply* coefficient rather than an input-*output* coefficient), which we will denote by $b_{i,j}$:

$$b_{i,j}^E = \frac{Z_{i,j}}{s_j} = \frac{Z_{i,j}}{x_j}\frac{x_j}{s_j} = a_{i,j}^E\frac{x_j}{s_j} \tag{17}$$

Thus, there is a close relationship between $b_{i,j}^E$ and the technological coefficient $a_{i,j}^E$. The factor of proportionality is (x_j/s_j). This is the share of total supply of product j that is provided by domestic output x_j. If the country does not import product j, we have $x_j = s_j$. In this case, the two coefficients coincide. Whenever imports of product j occur, $b_{i,j}^E$ will be smaller than $a_{i,j}^E$. The difference between the two can be interpreted as an indicator of self-sufficiency or import dependence. It is clear, however, that $b_{i,j}^E$ cannot be interpreted as a technological coefficient. The only technological coefficient is $a_{i,j}^E$, which is equal to $a_{i,j}^A$.

5 Implications for regional input-output modellers

The very technical discussion of the previous sections has important implications for regional input-output modellers. The reason is that regional input-output models are often constructed on the basis of regionalisation methods that adjust the national input-output table to regional conditions by applying mechanistic rules. These methods are laid out in the following.[8]

A variety of methods is based on a popular concept of regional science, the location quotient (LQ), which is generally interpreted as an indicator of an industry's relative over- or underrepresentation within a region (compared with the national average). The LQ is computed by using data that happen to be available. A direct measure of an industry's relative importance, such as the share of its output in total regional output, would be preferable. However, if such data are not available (at a satisfactory level of disaggregation), researchers often resort to employment data. The LQ of industry j is then computed according to the following formula:

$$LQ_j = \frac{L_j^R/\bar{L}^R}{L_j^N/\bar{L}^N} \tag{18}$$

The LQ, as computed by (18), is thus equal to the share of industry j in regional employment divided by the share of industry j at the national level. If $LQ_j > 1$ (i.e. the employment share of industry j at the regional level is larger than at the national level), industry j is said to be overrepresented. Conversely, if $LQ_j < 1$, industry j is said to be underrepresented. This procedure of describing the economic structure of a region, in comparison with other regions or the national average, has been standard practise in regional science for a long time.

---

[8]For a more extensive explanation of these methods, see Miller and Blair (2009, pp. 349–359).

The LQ has found use in the input-output literature as a tool for constructing regional input-output tables when detailed data are not available for the region to be studied. The idea is to regionalise the national input-output table by applying the LQ as correction factor of the A matrix. In the seminal paper by Schaffer and Chu, this is formulated as follows: "A location quotient of less than one means that the region imports some of its needs of output i. A location quotient greater than one means that the region exports some of output i" (Schaffer and Chu 1969, p. 85). Following this line of reasoning, Schaffer and Chu then explain what to do when the LQ is above or below unity: "If $LQ_i \geq 1$, we set $a_{i,j} = A_{i,j}$, where $a_{i,j}$ is the regional production coefficient (defined as $x_{i,j}/x_j$) and $A_{i,j}$ is the national production coefficient ($X_{i,j}/X_j$)" (Schaffer and Chu 1969, p. 85).

In other words, this procedure boils down to assuming that, if $LQ_i \geq 1$, the "regional production coefficient", $a_{i,j}$, is equal to its national counterpart. In the other case, Schaffer and Chu propose the following procedure: "If $LQ_i < 1$, local production is assumed to be inadequate to supply local needs—no exports can be made and imports are necessary. The regional production coefficient in row i may now be computed as $a_{i,j} = LQ_i A_{i,j}$" (Schaffer and Chu 1969, p. 86). Thus, in this case the "regional production coefficient" will be smaller than its national counterpart.

This approach is called the "simple location quotient" (SLQ) method. Since it has a number of shortcomings, various alternatives have been proposed. For a survey, see Miller and Blair (2009, pp. 349–359). However, in this paper, the focus is not on the particular shortcomings of the SLQ method; it is on the structure of the input-output table and the proper interpretation of the interindustry transactions matrix.

It is important to realise that Schaffer and Chu illustrated their discussion by means of an input-output table corresponding to our variant B. In their table, m_1 must be interpreted as "intermediate products imported for use by industry 1". It must *not* be interpreted as "imported products of type 1" (that would be the correct interpretation for Variant E tables). Thus, the reasoning behind the SLQ method (and, as remains to be shown, that of all other LQ methods) is based on an input-output table of the SIOT Variant B layout. The purpose of the present paper is to show that this has important implications for those who work with other variants of the SIOT.

In the context of a SIOT Variant B table, the reasoning behind the LQ methods is perfectly valid. As an empirical example, let us consider the case of electricity generation from coal in the city-state of Hamburg (one of the sixteen federal states that make up the Federal Republic of Germany). If somebody wanted to construct a Type B RIOT for Hamburg by means of the SLQ method, he would compute LQ_{COAL} using employment data for Hamburg and find that $LQ_{COAL} = 0$ since there are no coal mines in the region. The standard SLQ method calls for setting all entries in the corresponding row of the regional A matrix equal to zero. This is perfectly reasonable because $\mathbf{A}^B$, after all, refers only to regionally produced commodities.

However, in the context of a SIOT Variant A or E, this is not the case. A "naive" application of the SLQ method would mean that all entries in the row corresponding to coal would contain only zeroes. The problem here is that the matrix $\mathbf{A}^E$, in contrast to $\mathbf{A}^B$, refers not only to regionally produced commodities; it refers to the intermediate use of products, including imported products. Setting the first row of $\mathbf{A}^E$ equal to zero would mean that none of the regional industries used any coal whatsoever.

That is not a sensible result, because there is in fact a large coal-fired power plant in Hamburg.[9] This naive approach might lead to the erroneous conclusion that changes in final demand have no impact on coal consumption (and this in turn might lead to an underestimation of the environmental damage caused by the change in final demand).

Furthermore, a distinction must be made between the matrix of technical input-output coefficients ($\mathbf{A}^A$ or $\mathbf{A}^E$) and the matrix of input-supply coefficients ($\mathbf{B}^E$). The column labelled "coal" in matrix $\mathbf{B}^E$ should contain only zeroes, because the coal industry, being absent in Hamburg, does not use any inputs, so all inputs are equal zero, but supply is larger than zero due to imports. The SLQ procedure does not ensure that the "coal" column of the $\mathbf{B}^E$ matrix contains only zeroes. Thus, we must conclude that the SLQ method is not well suited to regionalising a SIOT of Variant A or E.

What about other LQ methods? Let us consider the CILQ method, which is based on the following formula (again taken from Schaffer and Chu):

$$\text{CILQ}_{i,j} = \frac{x_i / X_i}{x_j / X_j} \tag{19}$$

where $x_i(x_j)$ is regional output of industry $i(j)$ and $X_i(X_j)$ is national output of industry $i(j)$. In words: $\text{CILQ}_{i,j}$ is equal to the region's share in total output of industry i divided by the region's share in total output of industry j. Compared with the SLQ, the CILQ has the advantage that it does not only consider the relative size of industry i but also sets this in relation with industry j. However, this does not make it immune to the problem identified above. Once again, if industry i is coal mining, and coal mining does not exist in the region to be studied, all cells in row i will be equal to zero. This is perfectly reasonable for SIOT Variant B, but it does not make sense for SIOT Variant A or E. Thus, the CILQ is subject to the same limitation as the SLQ—it is not applicable to tables of Variant A or E.

What, then, can be done if a regional SIOT of Variant A or E is required? Here it is argued that, for these SIOT variants, the preferable regionalisation method should be based on the commodity balance (CB) approach (also known as supply-demand pool approach). This approach is based on the following equation, which is true by definition:

$$x_i + m_i^E = r_i + f_i \tag{20}$$

Equation (10) states that the sum of regional production and imports (the supply pool) must be equal to the sum of intermediate use and final use (the demand pool). Final use, in turn, can be decomposed into regional final use (regional consumption and capital formation) and exports, which yields:

$$\mathbf{x} + \mathbf{m} = \mathbf{r} + \mathbf{d} + \mathbf{e} \tag{21}$$

Note that the vector $\mathbf{m}$ in (21) is the column vector of imports by product as displayed in the SIOT Variant A (Table 2). Its transpose is the row vector $\mathbf{m}^E$, as displayed in the

---

[9]The Federal Ministry of the Environment maintains a database of power plants in Germany which shows a coal-fired power plant with a generation capacity of 205 MW in Hamburg (Umweltbundesamt 2011). Naturally, the coal it uses comes from imports by sea.

SIOT Variant E (Table 4). It is very different from the vector of imported intermediate products $\mathbf{z}^m$, as displayed in the SIOT Variant B (Table 3).

It is assumed that regional output x_i and regional final use d_i can somehow be estimated or measured. The next task is to estimate r_i. In order to do this, the "equal technology assumption" is invoked. This means that each industry in the region is assumed to operate with the same technological coefficients as at the national level. The technological coefficients can be calculated from the national input-output table according to (12) or (16).[10] Then, the same equations can be used to compute the $\mathbf{Z}$ matrix for the RIOT. Taking the row sum of $\mathbf{Z}$ yields the vector of intermediate use $\mathbf{r}$. Thus, the remaining task is to compute estimates for e_i and m_i. To do this, the trade balance b_i has to be computed. It is defined as:

$$\mathbf{b} = \mathbf{e} - \mathbf{m} \tag{22}$$

Solving (22) for $\mathbf{e}$ or $\mathbf{m}$ and substituting it into (21) yields:

$$\mathbf{b} = \mathbf{x} - \mathbf{r} - \mathbf{d} \tag{23}$$

Since data or estimates of $\mathbf{x}$, $\mathbf{r}$, and $\mathbf{d}$ are available, $\mathbf{b}$ can be computed from (23). However, it is not possible to compute actual exports and imports; (23) allows us only to compute the commodity balance. If we want to compute actual exports and imports, additional assumptions are required. For SIOT Variant A, this step is not very important because the columns labelled "exports" and "imports" may simply be subsumed by a column labelled "net exports". This is a way of evading the problem rather than solving it. If the goal is to construct a SIOT Variant E, the problem cannot be circumvented.

A very simple solution that has often been applied relies on the assumption that, for each product, the region is either import-dependent or not. That is, whenever the trade balance is positive, imports are assumed to be zero, and if the trade balance is negative, exports are assumed to be zero. This assumption rules out the possibility of cross-hauling (the simultaneous exporting and importing of similar products) and has been heavily criticised (Boomsma and Oosterhaven 1992; Richardson 1985). A more advanced treatment is possible with the Cross-Hauling Adjusted Regionalisation Method (CHARM), which estimates the amount of cross-hauling based on the heterogeneity of products (Kronenberg 2009, 2010). Either way, some estimate of $\mathbf{e}$ and $\mathbf{m}$ will be produced, and a SIOT Variant E can be constructed.

The CB method solves the problem outlined above, as can be seen with respect to the coal mining example. By multiplying the technological coefficients derived from the national IOT with the regional production vector, the matrix $\mathbf{Z}$ will be correctly estimated (subject to the drawbacks of the "equal technology assumption", of course). That is, column 1 of the matrix $\mathbf{Z}$ will contain only zeroes, because industry 1 (coal mining) does not exist and hence does not use any intermediate inputs. The entries in row 1, by contrast, will not be forced to equal zero, so the coal use of other industries is correctly respected. Given that regional production of coal is zero, the application

---

[10]Note that technological coefficients as defined in the present paper cannot be computed from SIOT Variant B tables.

of (22) will yield a negative trade balance for coal, and this will be reflected in the first element of the import vector **m**. In a Variant A table, the required imports of coal will be recorded in row 1, column "imports", and in the Variant E table they will be recorded in row "imports", column 1. All of this is absolutely correct.

On the other hand, neither the CB method nor the CHARM extension should be applied to SIOT Variant B tables. Aside from producing false results, there is also a logical contradiction. Setting up a commodity balance is possible only if the vector **m** is known. In a SIOT Variant B there is no such vector, therefore it is not possible to apply a CB method to such a table without using additional information.

6 Concluding remarks

The present paper shows that LQ methods can be applied to national input-output tables of the SIOT Variant B type, but they should not be applied to tables of Variant A or E. Conversely, CB methods, including CHARM, can be applied to Variant A or E tables, but not to type B tables. Applying LQ or CB methods to tables for which they are not suitable may result in misleading and implausible results.

The discussion of the previous sections is based purely on definitions and logical considerations. However, the following implications of the differences between the SIOT variants for regional input-output models have not yet been fully acknowledged in the literature on regional input-output modelling. The reason for this might be that much of that literature was written by authors based in the United States, where SIOT Variant B appears to be much more common than the other variants. Being used to working with this particular version of the SIOT, these authors did not devote much attention to the possible complications of working with other SIOT variants. This explains why neither Schaffer and Chu (1969) nor Miller and Blair (2009) discuss this problem.

Furthermore, in regional economics, it often makes sense to use the SIOT Variant B, as researchers are mostly interested in the effects of a certain final demand impulse on output, value added, and employment within the region of study. These questions coincide with the concerns of regional policy makers. When they decide, for example, on a particular investment project, they want to know how many jobs will be generated in the region for which they are responsible. However, input-output models are increasingly being used in the fields of environmental and ecological economics (Los 2011). Researchers from those fields are mostly interested in energy use, material consumption, greenhouse gas emissions and so on. For such research topics, SIOT Variant A or E is more useful. If we are concerned about climate change, we want to know the impact of final demand on, for example, electricity consumption, because electricity production is closely associated with greenhouse gas emissions. It does not matter much whether the electricity is produced in region A or region B (unless the electricity generation mix differs significantly). Therefore, we want to know how much electricity is actually used, and not whether it comes from domestic production or import. This is why ecological economists tend to prefer the allocation of imports according to Variant E. In the past, most of these studies have been undertaken for national economies. More recently, however, there seems to be a growing interest in

sustainable development and environmental policy at the regional level. As more and more researchers conduct environmental impact studies for individual regions, they will need to construct regional input-output models. Therefore, they need to be aware of the complications that stem from the different allocation of imports.

Appendix: Table 5

Table 5 Mathematical notation

$\mathbf{e}_i$	Exported products of type i
$\mathbf{s}_j$	Total supply of products of type j
$\mathbf{u}_i$	Total use of products of type i
$\mathbf{u}_i^d$	Total use of domestically produced products of type i
$\mathbf{u}_i^m$	Total use of imported products of type i
$\mathbf{v}_j$	Value added in industry j
$\mathbf{x}_j$	Total output of industry j
$\mathbf{y}_i^d$	Domestic final use of domestically produced commodities of type i
$\mathbf{y}_i^m$	Domestic final use of imported commodities of type i
$\mathbf{Z}_{i,j}^d$	Intermediate use of domestically produced products of type i in industry j
$\mathbf{Z}_{i,j}^m$	Intermediate use of imported products of type i in industry j

References

Bonfiglio A, Chelli F (2008) Assessing the behaviour of non-survey methods for constructing regional input–output tables through a Monte Carlo simulation. Econ Syst Res 20(3):243–258

Boomsma P, Oosterhaven J (1992) A double entry method for the construction of biregional input-output tables. J Reg Sci 32:269–284

Flegg AT, Tohmo T (2011) Regional input–output tables and the FLQ formula: a case study of Finland, Reg Stud. doi:10.1080/00343404.2011.592138

Flegg AT, Webber CD (1997) On the appropriate use of location quotients in generating regional input-output tables: reply. Reg Stud 31(8):795–805

Flegg AT, Webber CD (2000) Regional size, regional specialization and the FLQ formula. Reg Stud 34:563–569

Flegg AT, Webber CD, Elliott MV (1995) On the appropriate use of location quotients in generating regional input-output tables. Reg Stud 29(6):547–561

Holub HW, Schnabl H (1994) Input-Output-Rechnung: Input-Output-Analyse. München, Oldenbourg

Kronenberg T (2009) Construction of regional input-output tables using nonsurvey methods: the role of cross-hauling. Int Reg Sci Rev 32(1):40–64

Kronenberg T (2010) Erstellung einer Input-Output-Tabelle für Mecklenburg-Vorpommern. AStA Wirtsch Sozialstat Arch 4(3):223–248

Los B (2011) The output of input-output analysis: a bibliometric study (1996–2010). In: 19th international input-output conference, Alexandria, VA, USA

Miller RE, Blair PD (2009) Input-Output Analysis: Foundations and Extensions, 2nd edn. Cambridge University Press, Cambridge

Morrison WI, Smith P (1974) Nonsurvey input-output techniques at the small area level: an evaluation. J Reg Sci 14(1):1–14

Richardson HW (1985) Input-output and economic base multipliers: looking backward and forward. J Reg Sci 25:607–661

Schaffer WA, Chu K (1969) Nonsurvey techniques for constructing regional interindustry models. Pap Reg Sci 23:83–101

Stäglin R (2001) A step by step procedure to regionalized input-output analysis. In: Pfähler W (ed) Regional input-output analysis. Baden-Baden, Nomos

Tohmo T (2004) New developments in the use of location quotients to estimate regional input-output coefficients and multipliers. Reg Stud 38(1):43–54

Umweltbundesamt (2011) Datenbank "Kraftwerke in Deutschland". http://www.umweltbundesamt.de/energie/archiv/kraftwerke_in_deutschland.pdf. Accessed 12 January 2012

United Nations (1973) Input-output tables and analysis. New York