On the Construction of Regional Input-Output Tables with Imported Products inside the Transactions Matrix

Tobias Kronenberg[[1]](#footnote-1) & Johannes Többen

Research Centre Jülich, Institute of Energy and Climate Research

# Abstract

A number of recent papers discuss the accuracy of nonsurvey methods for the construction of regional input output tables (RIOTs). All these papers focus on RIOTs in which imports are allocated in a row in the South-West quadrant of the table (Type B tables). This means that, for example, the table shows how much the electricity sector spent on imported intermediate inputs, but it does not show how much of this was spent on coal, gas and oil. For certain applications, RIOTs of a different kind are preferable, with imported products included in the interindustry transactions matrix (Type Aor E tables). These tables show precisely the value of imported coal, oil and gas. So far, the literature focusses almost exclusively on Type B tables and offers little advice on the construction of Type A or E RIOTs.

We aim at closing this gap by, first, highlighting the conceptual differences between the two types of RIOTs and, second, illustrating them by using the example of Baden-Württemberg, one of Germany’s federal states. The statistical office of Baden-Württemberg produced type B *and*A tables for 1990. From the federal statistical office received corresponding tables for Germany as a whole. We use a variety of nonsurvey methods (SLQ, FLQ, CB, CHARM) to derive RIOTs of both types on the basis of the national tables and compare the derived tables to the official ones. Our findings suggest that LQ methods are better suited for deriving Type B tables and CB methods including CHARM are better suited for deriving Type A or E tables.

# Introduction

Since the publication of the first overview about non-survey methods for the regionalization of national input-output tables (NIOT) by Schaffer and Chu ([1969](#_ENREF_12)), a growing variety of both non-survey methods and applications of regional input-output tables, such as regional environmental impact studies can be observed in the literature. This variety of pure non-survey methods ranges from the members of the family of location-quotient methods like SLQ, FLQ or AFLQ to the commodity-balance method and its extensions such as CHARM ([Kronenberg, 2009](#_ENREF_6)). Naturally, the growing variety of methods induced a large and growing stream of literature discussing the strengths and weaknesses of these. ([Bonfiglio, 2009](#_ENREF_1); [Bonfiglio and Chelli, 2008](#_ENREF_2); [Flegg and Tohmo, 2011](#_ENREF_3); [Morrison and Smith, 1974](#_ENREF_9); [Richardson, 1985](#_ENREF_10); [Tohmo, 2004](#_ENREF_14)). However, our aim is to analyze strengths and weaknesses of non-survey from a different perspective: We argue that the choice of method should depend on the variant of import allocation of the symmetric input-output table that is to be regionalized.

The United Nations handbook on input-output analysis identified four different variants, labelled alphabetically from “A” to “D” ([United Nations, 1973](#_ENREF_15)). 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. As these are the most common variants, this paper concentrates on variant B and A, or E respectively.

In this paper we argue that LQ methods are suitable for type B tables, whereas CB methods are to be preferred for type E tables. This argument was laid out in theoretical terms in an earlier publication ([Kronenberg, in press](#_ENREF_7)). The purpose of the present paper is to illustrate the theoretical argument – and to test its validity – by means of an empirical study for the case of Baden-Württemberg, one of Germany’s federal states for which benchmarks in form of type B and type A tables for 1990 are available from the statistical office of Baden-Württemberg. For this comparison different non-survey methods, namely SLQ, CILQ, (A)FLQ, CB and CHARM, are applied to corresponding tables for Germany as whole.

The paper proceeds as follows: Section 2 explains the conceptual differences between type B, A and E tables and describes the interpretation of coefficients that arises from the differences. Section 3 identifies implications for the construction of RIOTs on a theoretical basis. Section 4 describes the empirical application to the case of Baden-Württemberg and presents the results. Section 5 draws a conclusion.

# Conceptual differences and interpretation of coefficients

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

Table 1: Basic data

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | |  |  |  |  |  |  |
|  |  |  |
|  |  |  |  |  |  |  |  |
|  | |  |  |  |  |  |  |
|  |  |  |
|  |  |  |  |  |  |  |  |
|  | |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  | |  |  |  |  |  |  |

Source: Kronenberg ([in press](#_ENREF_7)) IEK-STE 2012

The following conventions will be used: The subscript istands for products or commodities, the subscript *j* stands for industries or homogeneous branches. 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,  is element *i*, *j* of matrix . It reports the total amount of product *i* used by industry *j*. The amount of this which originates from domestic production is , and the amount which was imported is . Naturally, .

In addition to and , the basic data table contains the following elements. The vector  reports value added (i.e. primary inputs) by industry, and the vector  reports output (i.e. production) by industry. The vectors  and  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  reports exports of domestically produced commodities, whereas  reports exports of imported commodities (i.e. re-exports). It should be noted that re-exports are normally not included in input-output tables, so  will usually contain only zeroes. Finally, the vectors  and  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:

 (1)

where *n* is the number of products. Naturally, this relationship also holds for only domestically produced products or imported products:

 (2.A)

 (2.B)

These basic data provide the framework to discuss the conceptual differences of types of import allocation. Table 2 reports what we will call the SIOT Variant A[[2]](#footnote-2). At the core of the SIOT Variant A is an interindustry transactions matrix **Z**, which reports the entire intermediate consumption of products (domestically produced and imported). Mathematically, . Taking column sums of this matrix yields total intermediate consumption by industry, denoted by **z**. Taking row sums yield total intermediate consumption by product, denoted by **r**. By definition, summing **z** over *j* must yield the same result as summing **r** over *i*, so  (but  will usually not be true; it is possible but extremely unlikely).

Table 2: Symmetric input-output table, variant ‘A’

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Homogeneous branches | | | | Final uses | | Imports | Output |
| Products | 1 | … | *n* | Total | Domestic | Exports |
| 1 |  | … |  |  |  |  |  |  |
| … | … | … | … | … | … | … | … | … |
| *n* |  | … |  |  |  |  |  |  |
| Total interm. use / final use |  | … |  |  |  |  |  |  |
| Value added |  | … |  |  |  |  |  |  |
| Output |  | … |  |  |  |  |  |  |

Source: Kronenberg ([in press](#_ENREF_7)) IEK-STE 2012

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

 (3)

The bottom row of the SIOT variant A reports output by industry *xj*. Total output is the sum of *xj* over all *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:

 (4)

Rearranging terms yields:

 (5)

Going through row *i* of the SIOT variant A means going through equation (5). This is why imports are entered with a negative sign in the table.

The symmetry of SIOT variant A is captured by the following condition:

if . (6)

The Coefficients derived from variant A SIOTs have the following form:

 (7)

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*. Type A tables are suitable for environmental economic analysis, e.g. the estimation of resource consumption and associated harms to the environment, but they are not suitable to perform impact analysis. As for example imports of coal by electricity industry are recorded as a delivery of the coal industry to the electricity industry lumped together with coal form domestic production, it is not possible to say how much coal was imported and, thus, what the effects on domestic production are.

Table 3: Symmetric input-output table, variant ‘B’

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Homogeneous branches | | | | Final uses | | Output |
| Products | 1 | … | *n* | Total | Domestic | Exports |
| 1 |  | … |  |  |  |  |  |
| … | … | … | … | … | … | … | … |
| *n* |  | … |  |  |  |  |  |
| Imported products |  | … |  |  |  |  |  |
| Total interm. use / final use |  | … |  |  |  |  |  |
| Value added |  | … |  |  |  |  |  |
| Output |  | … |  |  |  |  |  |

Source: Kronenberg ([in press](#_ENREF_7)) IEK-STE 2012

Table 3 shows the symmetric input-output table, variant B[[3]](#footnote-3). 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 *mi*’s. In variant B, by contrast, imports are allocated by use, i.e. homogeneous branches and final users. denotes the column sums of matrix **Zm**. Thus,  is the value of imported products which were used as intermediate inputs by industry *j*. Accordingly,  denotes the value of total imports used as intermediate inputs.  denotes the value of products that were consumed by final users in the country, and  denotes the value of imported products used for exports (re-exports).  denotes the total value of imports.

It is important to realize that **zm** is very different from **m**. The latter is a column vector of length *n* (the number of different products), and element *mi* is interpreted as “imported products of type *i*”. The former is a row vector of length *n*, and element  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, but vector **zm** contains only those products imported for intermediate, as imported products for final use are recorded elsewhere.

The type B coefficients are then defined as:

 (8)

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’s define a trading coefficient *ti,j* as:

 (9)

In words, *ti,j* is the share of input *i* used by industry *j* that originates from domestic production. Conversely,  can be interpreted as the import share. Using equations 7, 8, and 9, we can write the relationship between  and  as:

 (10)

Thus,  will generally be smaller than . The  coefficients differ from the true technological coefficients () due to international trade. Therefore, they cannot not be interpreted as technological coefficients. They are a mixture of technology and trade.

As, for example, imported coal of the electricity industry is lumped together all other imported goods, it is impossible to say how much coal was actually used in the production of electricity. But it is possible how much coal from domestic sources was used, so that type B tables are useful for estimating the impact of final demand on national/regional employment, value added etc.

Table 4: Symmetric input-output table, variant ‘E’

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Homogeneous branches | | | | Final uses | | | Total use |
| Products | 1 | … | *n* | Total | Domestic | Exports | Total |
| 1 |  | … |  |  |  |  |  |  |
| … | … | … | … | … | … | … | … | … |
| *n* |  | … |  |  |  |  |  |  |
| Total interm. use / final use |  | … |  |  |  |  |  |  |
| Value added |  | … |  |  |  |  |  |  |
| Output |  | … |  |  |  |  |  |  |
| Imports of similar goods |  | … |  |  |  |  |  |  |
| Total supply |  | … |  |  |  |  |  |  |

Source: Kronenberg ([in press](#_ENREF_7)) IEK-STE 2012

Table 4 shows how Eurostat currently compiles its input-output tables according to the ESA 95 guidelines. This variant is called variant ‘E’ for ‘Eurostat’.

The total supply of a commodity is equal to domestic production plus imports of similar commodities:

 (11)

The row vector **mE** refers to imports by commodity. **mE** stands for ‘vector of imports constructed the Eurostat way’. 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, **mE** is the transpose of **mA**.

At the core of Variant E is the interindustry transactions matrix **Z**, as in variant A. Taking column sums of this matrix yields total intermediate consumption by industry, denoted by **z**. Taking row sums yield total intermediate consumption by product, denoted by **r**. By definition, summing **z** over *j* must yield the same result as summing **r** over *i*, so  (but  will usually not be true; it is possible but extremely unlikely).

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

 (12)

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

 (13)

Finally, it is true by definition that

 (14)

Thus, the IOT is symmetric in the sense that when .

Finally, which coefficients do we have in variant E? This depends on what we put in the denominator – output or supply. If we put output in the denominator, we are performing exactly the same calculation as in (7). Mathematically:

 (15)

Thus, variant E also allows the computation of technological coefficients.

If we divide *Zi,j* by *sj*, we get a different kind of coefficients, which we will call *bi,j*:

 (16)

Thus, there is a close relationship between  and the technological coefficient . The factor of proportionality is . This is the share of total supply of product *j* which is provided by domestic output *xj*. If the country does not import product *j*, we have . In this case, the two coefficients coincide. Whenever imports of product *j* are larger than zero,  will be smaller than . The difference between the two can be interpreted as an indicator of self-sufficiency or import dependence. It is clear, however, that  cannot be interpreted as a technological coefficient. The only technological coefficient is , which is equal to .In the case of variant E SIOTs it is, for example, possible to how much coal was used by providers of electricity and it is also possible how much coal was imported, which makes these tables suitable to to estimate the effects of final demand on resource use **and** domestic/regional employment, value added etc.

# Implications for non-survey methods

The very technical discussion of the previous sections has important implications for regional input-output modellers. The reason for this 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[[4]](#footnote-4).

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 to the national average). The LQ is computed by using data that happens to be available. Preferable is a direct measure of an industry’s relative importance, such as the share of its output in total regional output. 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 following formula:

 (17)

In words, the LQ as computed by equation (17) is equal to the share of industry *j* in regional employment divided by the share of industry *j* at the national level. If  (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  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 practice in regional science for a long time.

The LQ has found use in the input-output literature as a tool for constructing regional input-output tables when detailed data is 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](#_ENREF_12)). Following this line of reasoning, Schaffer and Chu then explain what to do when the location quotient is greater or smaller than one: “If , we set , where  is the regional production coefficient (defined as ) and  is the national production coefficient (). Knowing regional industry outputs, , and having established , we may easily compute regional interindustry flows” ([Schaffer and Chu, 1969, p. 85](#_ENREF_12)). They then propose the following formula for computing :

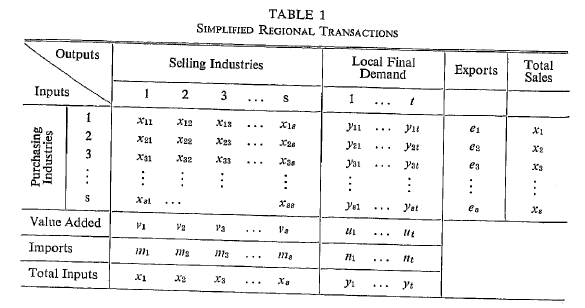
. (18)

In other words, this procedure boils down to assuming that if  the “regional production coefficient”, , is equal to its national counterpart. In the other case, Schaffer and Chu propose the following procedure: “If , 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 ” ([Schaffer and Chu, 1969, p. 86](#_ENREF_12)). 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 shortcomings, various alternatives have been proposed. For a survey, see Miller and Blair ([2009, pp. 349-359](#_ENREF_8)). 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.

The input-output table used by Schaffer and Chu looks like this:

Table 5: the input-output table of Schaffer and Chu



Source: ([Schaffer and Chu, 1969](#_ENREF_12)) IEK-STE 2012

Obviously, this table is of the SIOT Variant B format. This means that, for example,  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 convince you, the reader, 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. Let us assume an extreme case: Industry 1 is coal, and the regional economy happens to have no coal mines at all. Consequently, . When equation (18) is applied, all entries in row 1 will be zero. If the other industries in the region need coal as an intermediate input, they will import coal, and these imports will be recorded in the row labelled “imports”, but not in the X matrix. Thus, the reasoning of Schaffer and Chu is perfectly valid.

However, in the context of a SIOT Variant A or E this is not the case. Let us consider again Table 5, the SIOT Variant E. The application of equation (18) would mean that the first row of the Z matrix contains only zeroes. But the Z matrix in Table 5, unlike the X matrix of Schaffer and Chu, refers not only to intraregional transactions; it refers to the intermediate use of products, including imported products. Setting the first row of the Z matrix equal to zero would mean that none of the regional industries use any coal whatsoever. What’s worse is that the SLQ procedure does not even ensure that the first column of the Z matrix contains only zeroes (but it should, for if industry 1 is not present in the region it does not use any inputs and consequently the first row of the Z matrix must contain only zeroes). This example shows that the SLQ method is not well-suited for regionalising SIOT of Variant A or E.

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

 (19)

The CILQ method 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 method is subject to the same limitation as SLQ – it is not applicable for tables of SIOT Variant A or E. The same is true for recent LQ vairants such as FLQ and AFLQ , as these methods are based on multiplicative adjustments CILQ (and SLQ for elements on the main diagonal) and, therefroe, do not solve the problem for variant A and E tables outlined above.

What, then, can be done if a regional SIOT of Variant A or E is required? The present paper argues 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:

. (20)

In words, equation (20) 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:

. (21)

Note that the vector **m** which appears in (21) is the column vector of imports by product as displayed in the SIOT Variant A (Table 3). Its transpose is the row vector **mE** as displayed in the SIOT Variant E (Table 5). It is very different from the vector of imported intermediate products **zm** as displayed in the SIOT Variant B (Table 4).

It is assumed that regional output  and regional final use  can somehow be estimated or measured. The next task is to estimate . 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 on the national level. The technological coefficients can be calculated from the national input-output table according to (7) or (15)[[5]](#footnote-5). Then, the same equations can be used to compute the Z matrix for the RIOT. Taking the row sum of **Z** yields the vector of intermediate use **r**. Thus, the remaining task is to compute estimates for  and . To do this, the trade balance  has to computed. It is defined as:

 (22)

Solving (22) for **e** or **m** and substituting it into (21) yields:

 (23)

Since data or estimates of **x**, **r**, and **d** are available, **b** can be computed from (23). However, it is not possible to compute actual exports and imports; (23) allows us only to compute the eponymous 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 but not solving it. If the goal is to construct a SIOT Variant E, this 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 ([Richardson, 1985](#_ENREF_10))). 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](#_ENREF_6))). Either way, some estimate of **e** and **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 **Z** will be correctly estimated (subject to the drawbacks of the “equal technology assumption”, of course). That is, column 1 of the matrix **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 of (23) 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 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.

# Empirical application to Baden-Württemberg

The aim of this chapter is to compare the performance of various non-survey methods in constructing regional type B and type E tables for Baden-Württemberg. Regional type B tables are constructed by making use of several LQ-methods, namely SLQ, FLQ, AFLQ, and CB, whereas for type E tables CHARM is used instead of the ordinary CB along with LQ-Methods. The performance is measured be comparing estimated coefficients and output multipliers with those that were calculated from benchmark tables provided by the statistical office of Baden-Württemberg. National type B and E SIOTs were provided by the federal statistical office. For the regionalization we use data on regional output by industry from the regional benchmark tables for all methods. As the construction of variant E tables requires an estimation of gross imports by industry, we state the common assumption of import-dependency for LQ methods as mentioned above. CHARM does not require stating such assumptions. The performance of non-survey methods to estimate coefficients is measured by **M**ean **P**ercentage **A**verage **D**eviation (MAPE), which is defined as:

(24)

Where denotes coefficients calculated from the estimated SIOT of variant B (tilde generally indicates coefficients and multipliers calculated from benchmark tables). The next two statistics compare estimated total intermediate inputs by industry with those calculated from the benchmark tables. The first one is the **W**eighted **D**eviation (WD) and the second the **W**eighted **A**bsolut **D**eviation (WAD) with industry shares in output as weighting factor:

(25)

(26)

The measurements (24), (25) and (26) are also applied to the input-supply coefficients derived from type E tables. The measurement of deviations in output multipliers is also done by applying (25) to column sums of the Leontief inverse .

In the case of FLQ and AFLQ, performance crucially depends on the value of the parameter , which adjusts for for regional size and varies over regions. For this reason we distinguish two cases: In the first case we use values that were suggested in literature, namely for FLQ and for AFLQ ([Flegg and Webber, 2000](#_ENREF_4); [Tohmo, 2004](#_ENREF_14)). In the second case the value for is chosen such that the WAD of the output multipliers is minimized. The results of FLQ and AFLQ as presented by the figures in the following subsection are based on an optimal .

## Type B tables

Table 5 presents the results of the performance of non-survey methods in regionalizing a variant B SIOT for Baden-Württemberg. It is observed that FLQ and AFLQ deliver the best result in terms of accuracy in the estimation of output multipliers, even if the value of is taken from the literature (in this case AFLQ performs slightly better than the original FLQ). The optimal values of in the case of Baden-Württemberg are 0.2 for FLQ and 0.28 for AFLQ respectively. With a WAD of 15.13% and 11.72% CB and SLQ perform clearly worse in estimating regional output-multipliers.

On the level of individual cells (MAPE) FLQ and AFLQ also perform better than CB and SLQ. Considering the results for output multipliers it is on the first sight somewhat surprising that FLQ(opt) and AFLQ(opt) exhibit greater deviations than the FLQ and AFLQ with the value of taken from the literature, but the corresponding WD indicates that the reason for it is a greater compensation of positive and negative errors on the level of individual cell and columns respectively. This effect also explains why FLQ generates the greatest errors in terms of estimating the share of total (domestic) intermediate inputs in output by industry (WAD), although it estimates output multipliers properly.

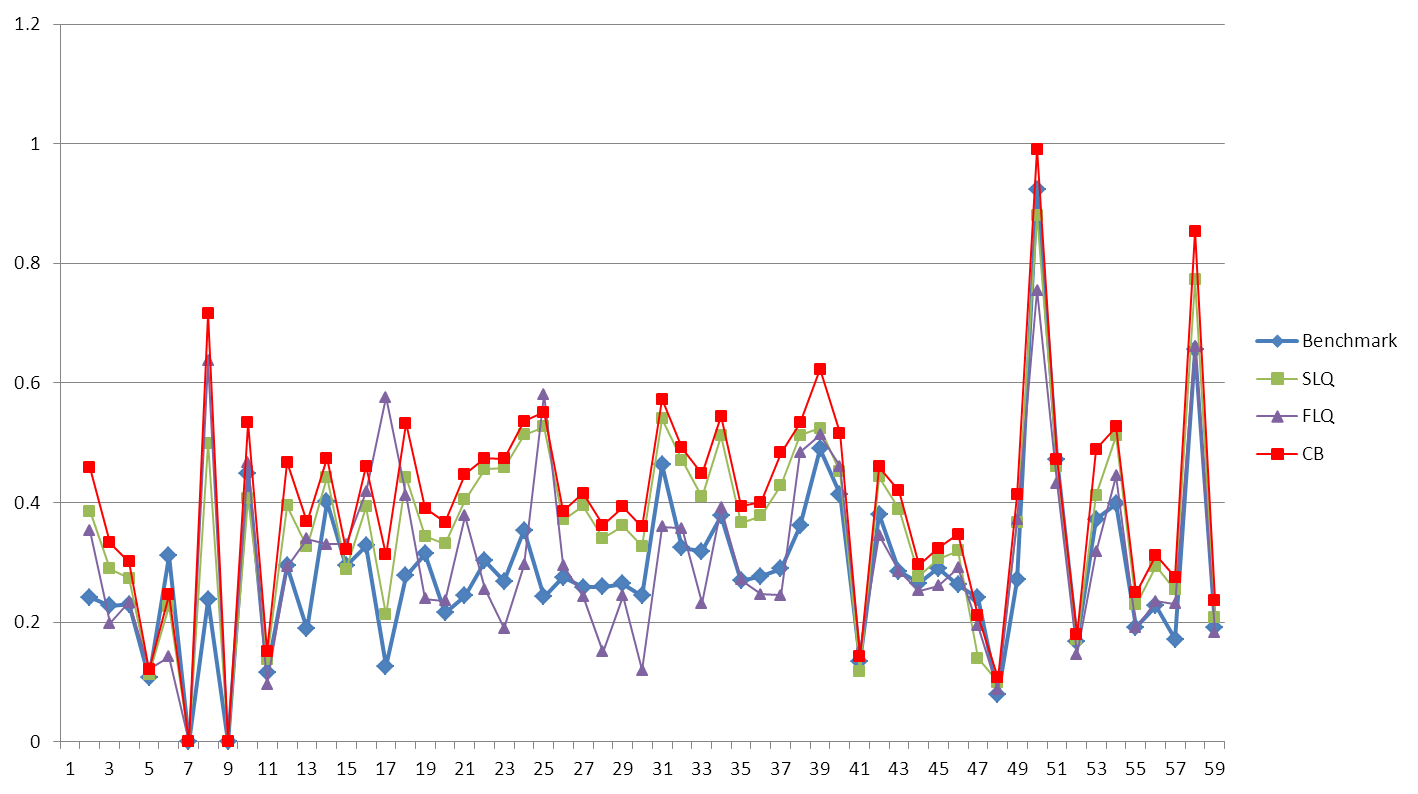
Table 5: Performance in regionalization of type B tables

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Coefficients** | | | **Outputmultipl.** |
| **Method** | **WD** | **WAD** | **MAPE** | **WAD** |
| **CB** | -39.99% | 40.11% | 58.43% | 15.13% |
| **SLQ** | -29.37% | 30.28% | 46.95% | 11.72% |
| **FLQ (opt.)** | 1.39% | 61.60% | 39.88% | 4.80% |
| **FLQ (0.3)** | 14.02% | 80.28% | 37.81% | 8.72% |
| **AFLQ (opt.)** | 1.53% | 13.18% | 40.49% | 5.03% |
| **AFLQ (0.4)** | 17.03% | 20.05% | 37.81% | 8.12% |

Source: authors’ calculations IEK-STE 2012

Figure 1 presents the shares of (domestic) intermediate inputs in output by sector. AFLQ(opt) is excluded for sake of clearness, as there is no big difference to the result of FLQ (opt). It can generally be said that in contrast to SLQ and CB, FLQ, despite from being closer to the benchmark table for most industries, does not overestimate intermediate inputs uniformly, but rather underestimates intermediate inputs in many cases.

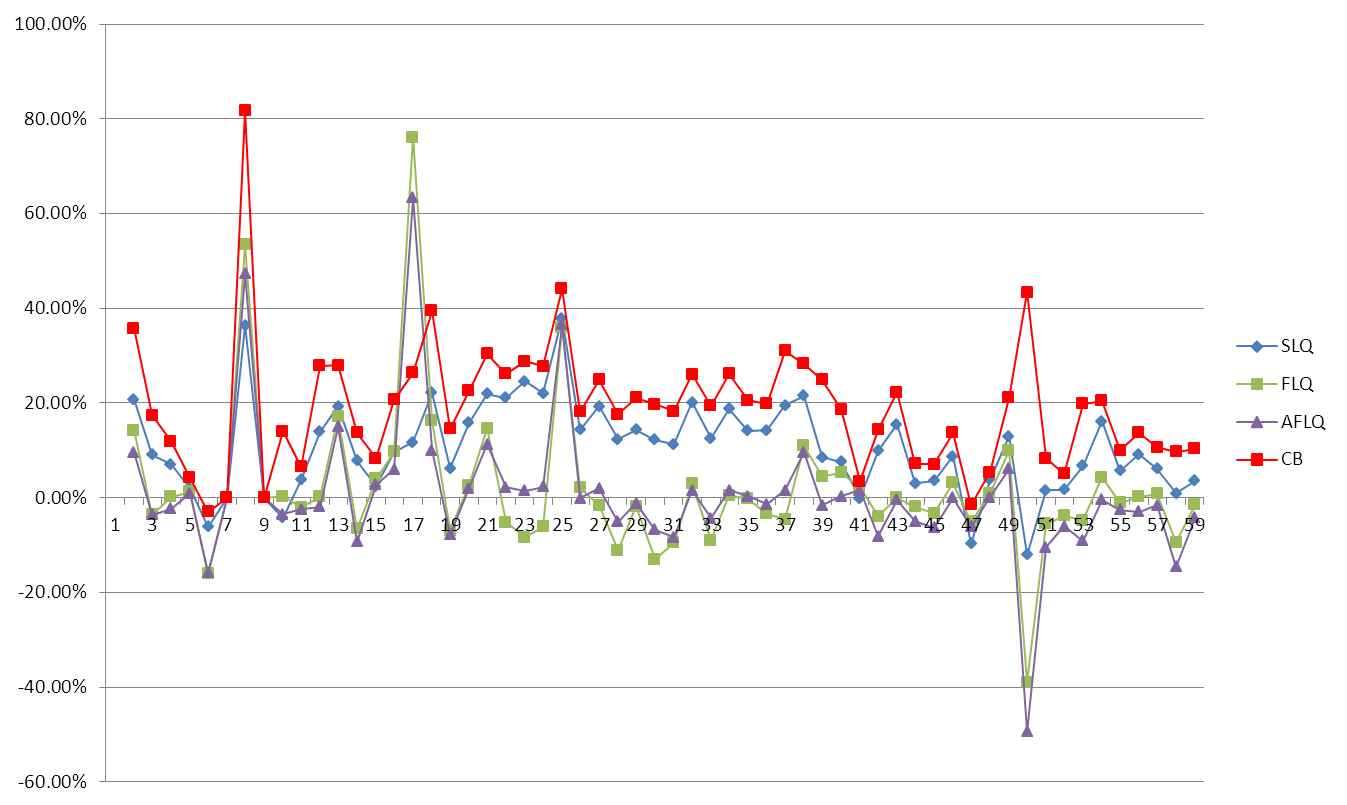
Figure 1: Share of (domestic) intermediate inputs in output by industry (%)



Source: authors’ calculations IEK-STE 2012

Figure 2 presents the deviations of estimated output multipliers from their benchmark counterparts on the level of industries. Whereas CB and SLQ generally overestimate output multipliers, despite of few cases, deviations of FLQ and AFLQ estimates are considerably smaller and have positive as well as negative directions. The pattern of overestimation over several industries of CB and SLQ show that both methods delivers results that go in a similar direction indicating the both close theoretical relationship of both methods ([Round, 1972](#_ENREF_11)), whereby the deviations of CB from the benchmark are generally greater.

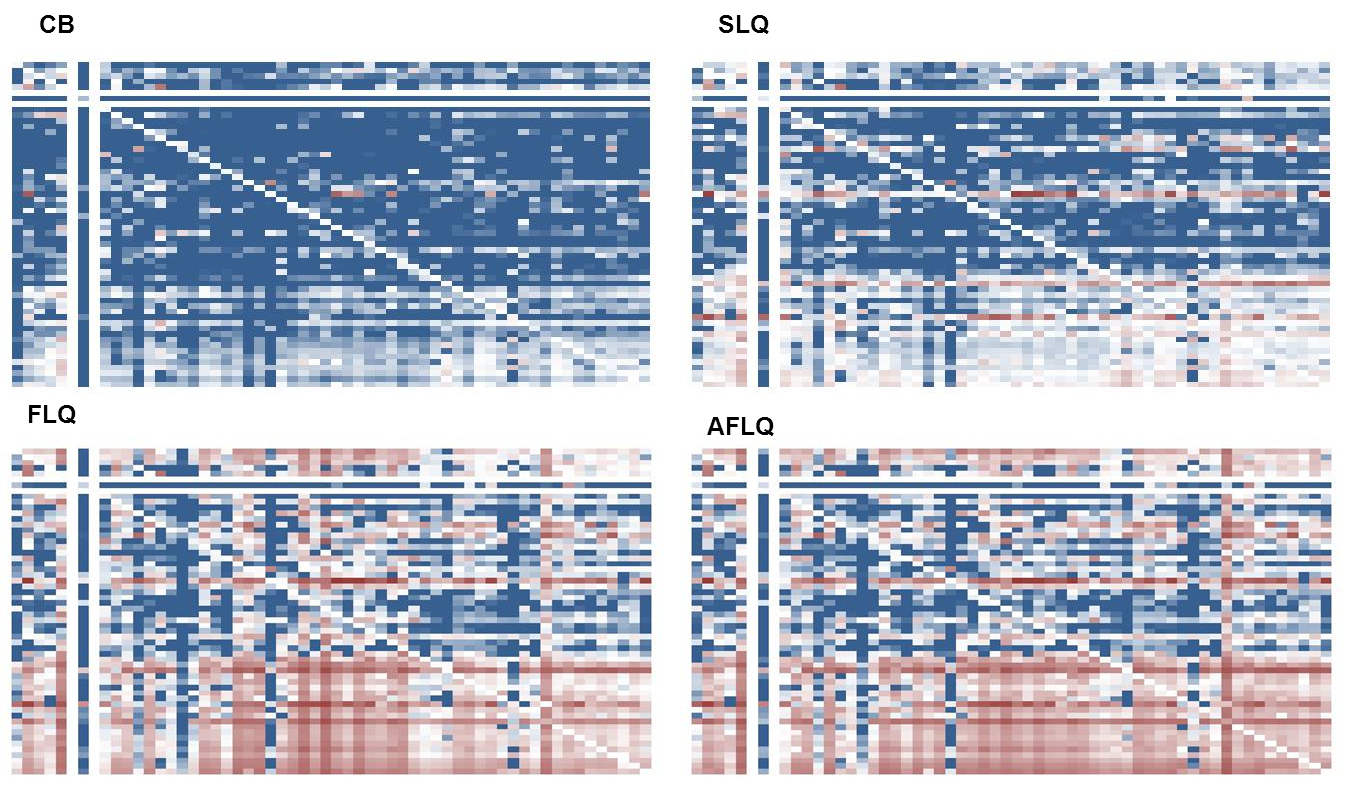
Figure 2 Deviation from benchmark multipliers by industry (%)



Source: authors’ calculations IEK-STE 2012

Figure 3 presents the deviation of the elements of the Leontief-inverse estimated by non-survey methods from those of the benchmark table. These deviations are indicated by colors running from dark red indicating an element that underestimated by 100% or more, to dark blue indicating an element that is overestimated by 100% and more. Whereas the commodity balance method strongly overestimates almost all elements, FLQ and AFLQ underestimate many elements at the bottom indicating purchases from service industries, which results in smaller deviations of output multipliers, due to compensation effects within a column. All of the methods have a tendency to overestimate especially the elements representing purchases from the manufacturing industry. All of the four methods estimated the nonexistence of coal mining and natural gas in Baden-Württemberg correctly indicated by the white stripes in left and at the top of the Leontief-inverse. The reason for the relatively good performance of all four methods in estimating output multipliers considering the strong deviations on the level of individual cells is that the diagonal elements are estimated properly.

Figure 3 Relative deviations of elements of the Leontief-Inverse



Source: authors’ calculations IEK-STE 2012

## Type E tables

Table 6 shows the results for non-survey methods for the construction of type E RIOTS. According to the WD and MAPE statistics the results of CHARM show that this method delivers the best result in terms of estimating coefficients and output multipliers. On an aggregate level ALFQ delivers the second best results followed by FLQ. SLQ performs worst. It has to be noted that the optimal value of for FLQ is -0,2, which is a clear contradiction, as the parameter is defined to have a value between zero and one. The WD measurement for CHARM and SLQ shows that these methods overestimate coefficients in total, whereas the results for FLQ and AFLQ indicate a tendency to underestimate them. The reason for it might be the fact, that FLQ and AFLQ are designed to estimate inputs from regional production and scales most coefficients down. That would mean applied to a type E, that a regional industry has a higher productivity, as these are technological coefficients.

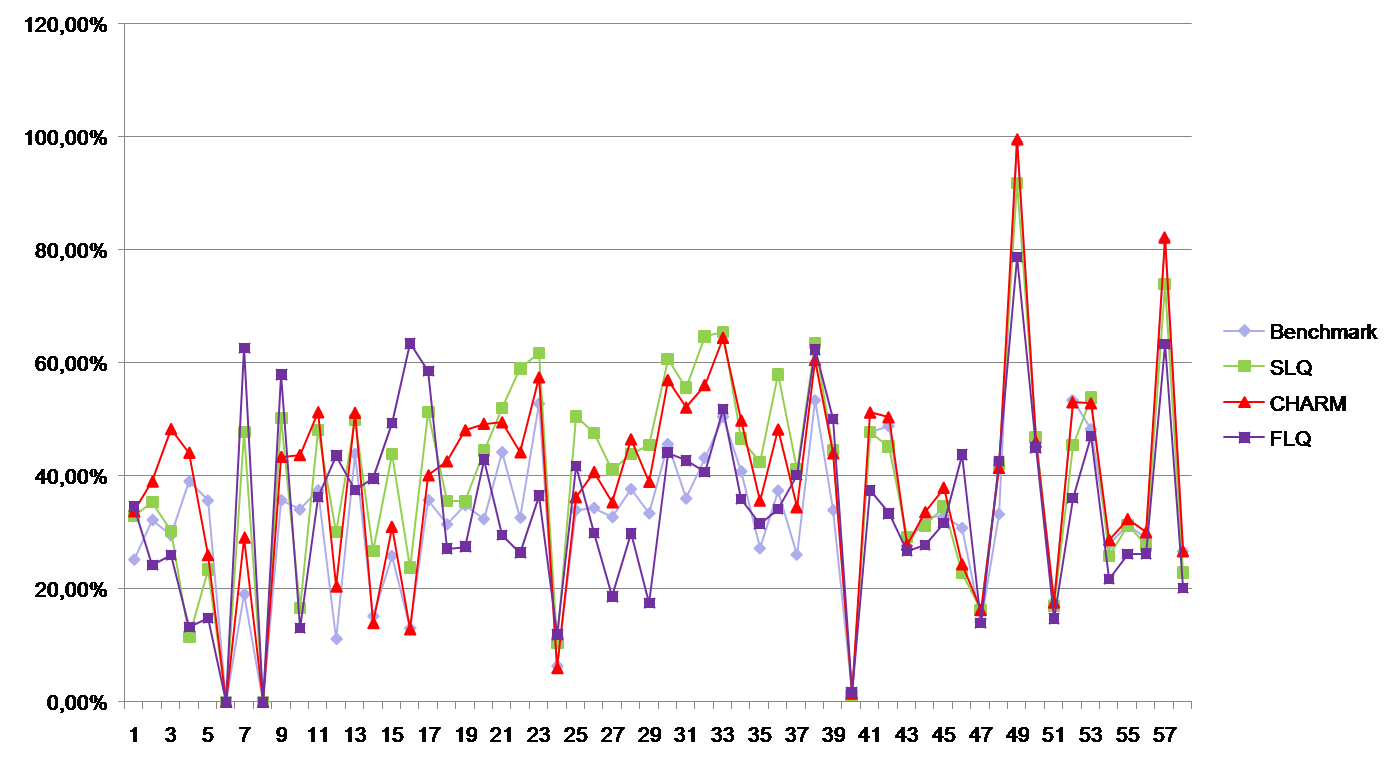
Table 6: Performance of non-survey methods in construction of type E tables

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Coefficients** | | | **Outputmultipl.** |
| **Methode** | **WD** | **WAD** | **MAPE** | **WAD** |
| **CHARM** | 12,53% | 12,87% | 22,36% | 8,21% |
| **SLQ** | 11,92% | 17,83% | 35,33% | 10,11% |
| **FLQ (opt.)** | -13,98% | 22,94% | 31,77% | 10,57% |
| **FLQ (0.3)** | -11,19% | 23,56% | 36,67% | 10,75% |
| **AFLQ (opt.)** | -7,63% | 19,17% | 36,07% | 8,56% |
| **AFLQ (0.4)** | -5,89% | 19,58% | 37,26% | 8,71% |

Source: authors’ calculations IEK-STE 2012

Figure 4 presents the shares of total intermediate inputs in output by industries of non-survey and benchmark tables. CHARM and SLQ overestimate total intermediate inputs of most industries, whereas FLQ estimates are more strongly scattered

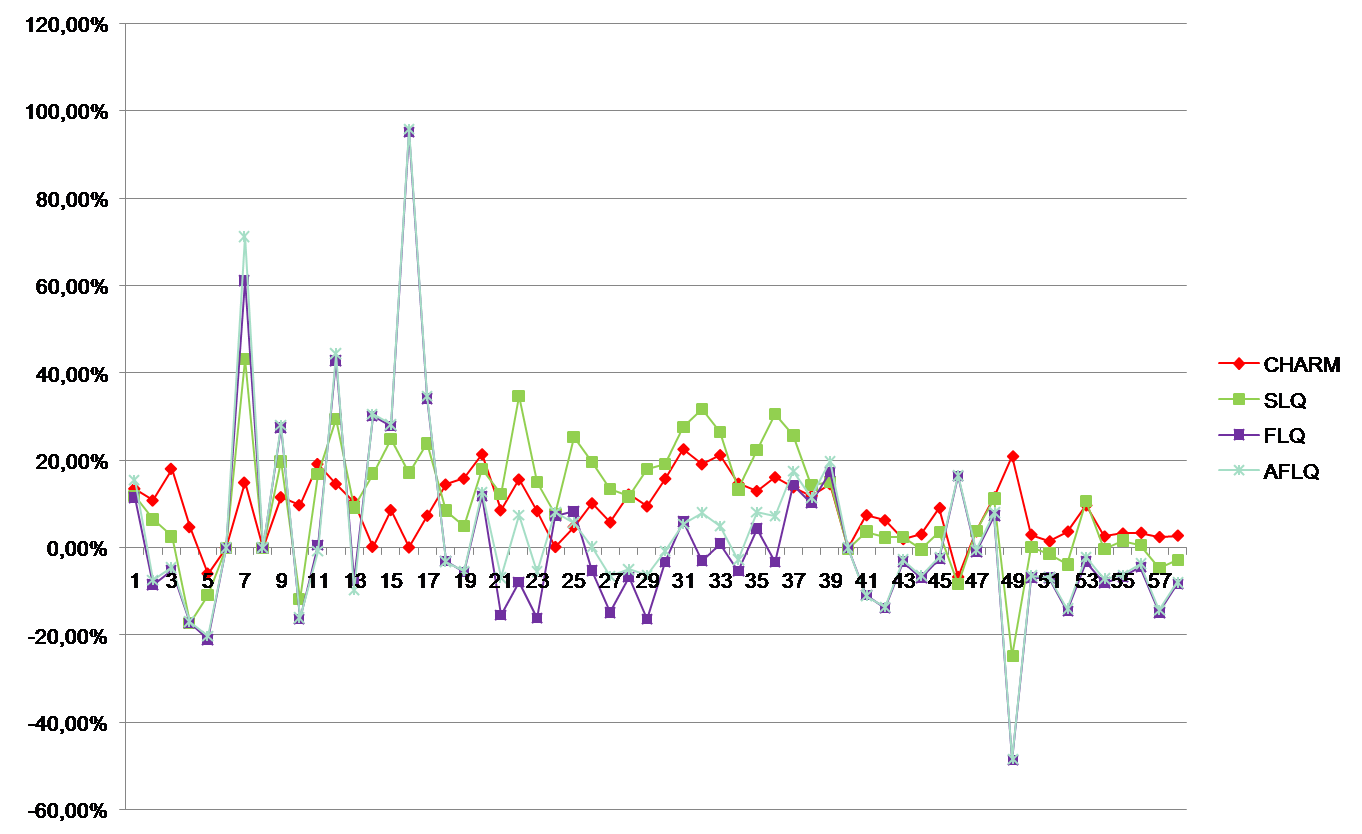
Figure 4: Share of (total) intermediate inputs in output by industry (%)



Source: authors’ calculations IEK-STE 2012

Figure 5 shows deviations of estimated output multipliers by industry from those derived from the benchmark RIOT. Despite water supply industry (5) and inland navigation (47), CHARM overestimates output multipliers of all industries. SLQ also overestimates most of the multipliers, whereas FLQ and AFLQ underestimate many multipliers. This concerns for example industries such as machinery (21) and manufacturing of motor vehicles (23), which are strongly overrepresented industries in Baden-Württemberg.

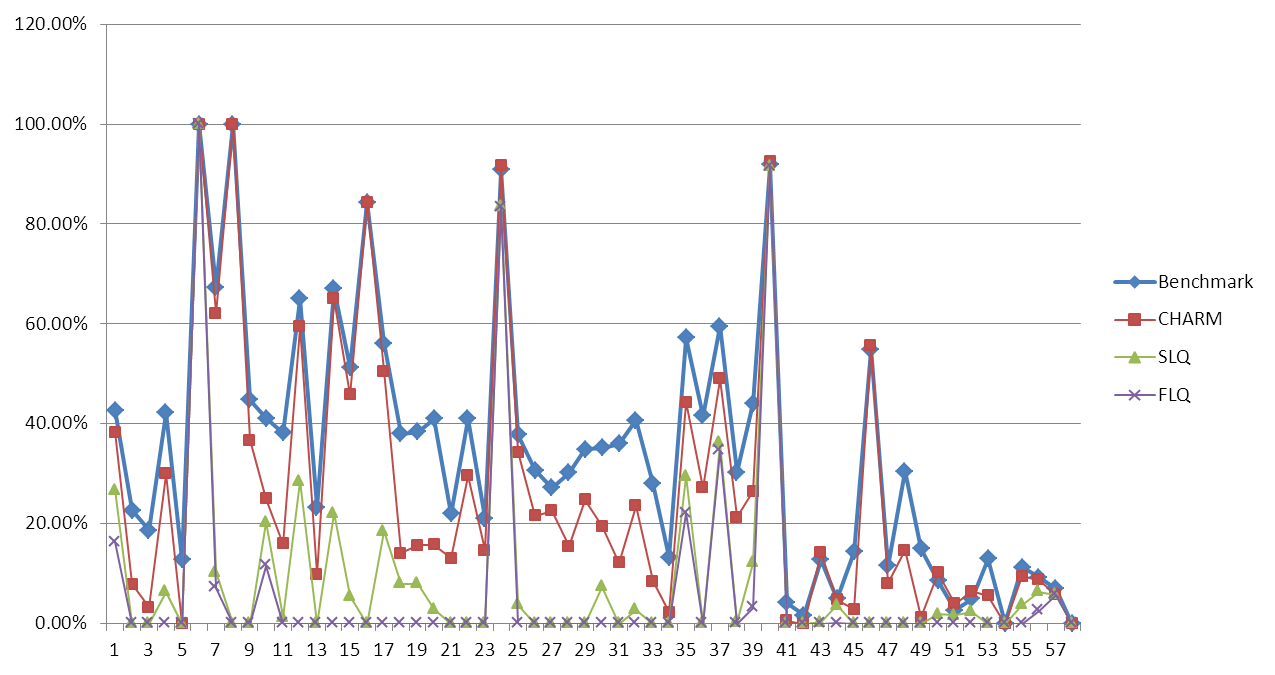
Figure 5: Deviation from benchmark multipliers by industry (%)



Source: authors’ calculations IEK-STE 2012

Figure 6 shows estimated import shares by industry in comparison to those of the benchmark RIOT. All of the methods underestimated imports of each industry. Whereas CHARM delivers relatively good estimates of imports, the results of SLQ are much weaker and FLQ performs even worse, as imports are set to zero, when a sector has a positive trade balance.

Figure 6: Import-quotas by industry (%)



Source: authors’ calculations IEK-STE 2012

Figure 7 shows the deviation of the elements of the Leontief inverse estimated by non-survey methods from those of the benchmark table using the same color scale as Figure 3. CHARM overestimates almost all cells, but the deviations are in most cases smaller than those of other methods. FLQ and AFLQ seem to have a tendency of underestimating many elements of the Leontief-inverse, when applied to type E tables. Also all of the LQ methods fail in the estimation of elements associated with purchases of coal and natural gas as discussed in chapter 3 (red lines at the top).

Figure 7: Relative deviations of elements of the Leontief-Inverse

|  |  |
| --- | --- |
| CHARM | SLQ |
| Variante E CHARM.jpg |  |
| FLQ | AFLQ |
|  |  |

Source: authors’ calculations IEK-STE 2012

## Conclusion

To sum up, the present paper shows that the choice of the regionalization method should depend on the type of SIOT to be regionalized. For the construction of type B tables, LQ methods are the correct choice (preferably FLQ or perhaps AFLQ). For the construction of type E tables, it appears that CB methods like CHARM are the correct choice. These findings lend empirical support to the theoretical argument outlined in an earlier publication ([Kronenberg, in press](#_ENREF_7)).

Nevertheless, we would argue that a pure nonsurvey approach is not sufficient for the construction of high-quality regional input-output tables. Even the more advanced methods may still lead to misleading results. For actual empirical work, perhaps for projects commissioned by public agencies or other relevant actors, it is important that the nonsurvey approach can provide only a first “draft” version of the RIOT. Additional data should be used to check the plausibility of the results, to ensure compatibility with official statistics, and to improve the accuracy of the implied multipliers.

# References

Bonfiglio, A. (2009). 'On the Parameterization of Techniques for Representing Regional Economic Structures', *Economic Systems Research*, vol. 21, pp. 115-127.

Bonfiglio, A. and F. Chelli (2008). 'Assessing the Behaviour of Non-Survey Methods for Constructing Regional Input–Output Tables through a Monte Carlo Simulation', *Economic Systems Research*, vol. 20(3), pp. 243-258.

Flegg, A. T. and T. Tohmo (2011). 'Regional Input–Output Tables and the FLQ Formula: A Case Study of Finland', Regional Studies, first published on 25 August 2011 (iFirst), doi: 10.1080/00343404.2011.592138.

Flegg, A. T. and C. D. Webber (2000). 'Regional Size, Regional Specialization and the FLQ Formula', *Regional Studies*, vol. 34, pp. 563-569.

Holub, H. W. and H. Schnabl (1994). *Input-Output-Rechnung: Input-Output-Analyse*, München, Wien: Oldenbourg.

Kronenberg, T. (2009). 'Construction of Regional Input-Output Tables Using Nonsurvey Methods: The Role of Cross-Hauling', *International Regional Science Review*, vol. 32(1), pp. 40-64.

Kronenberg, T. (in press). 'Regional input-output models and the treatment of imports in the European System of Accounts (ESA)', *Jahrbuch für Regionalwissenschaft*.

Miller, R. E. and P. D. Blair (2009). *Input-Output Analysis: foundations and extensions*, 2nd ed. Cambridge: Cambridge University Press.

Morrison, W. I. and P. Smith (1974). 'Nonsurvey Input-Output Techniques at the Small Area Level: An Evaluation', *Journal Of Regional Science*, vol. 14(1), pp. 1-14.

Richardson, H. W. (1985). 'Input-output and economic base multipliers: Looking backward and forward', *Journal Of Regional Science*, vol. 25, pp. 607-661.

Round, J. L. (1972). 'Regional Input-Output Models in the U.K.: A Reappraisal of Some Techniques', *Regional Studies*, vol. 6, pp. 1-9.

Schaffer, W. A. and K. Chu (1969). 'Nonsurvey Techniques for Constructing Regional Interindustry Models', *Papers of the Regional Science Association*, vol. 23, pp. 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', *Regional Studies*, vol. 38(1), pp. 43-54.

United Nations (1973). *Input-Output Tables and Analysis*, New York.

1. Corresponding author. E-Mail: [t.kronenberg@fz-juelich.de](mailto:t.kronenberg@fz-juelich.de). Phone: +49 2461 61 1721 [↑](#footnote-ref-1)
2. This is what Holub and Schnabl ([Holub and Schnabl, 1994](#_ENREF_5)) call „Variante A1“. At the regional level, it is closely related to what Stäglin([Stäglin, 2001](#_ENREF_13)) calls the “technological version”. [↑](#footnote-ref-2)
3. This is what Holub and Schnabl ([Holub and Schnabl, 1994](#_ENREF_5)) call „Variante B“ and Stäglin([Stäglin, 2001](#_ENREF_13)) calls “regional version”. [↑](#footnote-ref-3)
4. For a more extensive explanation of these methods, see Miller and Blair ([Miller and Blair, 2009, pp. 349-359](#_ENREF_8)). [↑](#footnote-ref-4)
5. Note that technological coefficients as defined in the present paper cannot be computed from SIOT Variant B tables. [↑](#footnote-ref-5)