STE Preprint

14
2007

Forschungszentrum Jülich
in der Helmholtz-Gemeinschaft

How Can Regionalization Methods Deal With Cross-Hauling?

Tobias Kronenberg

Institut für Energieforschung (IEF)
Systemforschung und Technologische Entwicklung (STE)

Contents

How Can Regionalization Methods Deal With Cross-Hauling?

Tobias Kronenberg[1)]

1) Research Centre Jülich, Institute of Energy Research - Systems Analysis and Technology Evaluation (STE), D-52425 Jülich, Germany

Abstract

The construction of regional input-output tables is expensive and time-consuming. Therefore, they are usually estimated using regionalization methods, which adjust the national IO-table to the region in question. However, common regionalization methods result in biased estimates of multipliers because they do not take cross-hauling into account. Moreover, they are not compatible with the accounting procedures of ESA 95 and cannot be applied to IO-tables with indirect allocation of imports.

This paper provides an interpretation of the literature on regionalization methods and presents a new approach based on an estimate of product heterogeneity, which addresses the problem of cross-hauling and is applicable to tables with indirect allocation of imports.

Keywords

regional input-output analysis, nonsurvey methods, hybrid methods, cross-hauling

Contribution to

16th International Input-Output Conference, Istanbul, 2-6 July 2007

I Introduction

Regional input-output tables are very useful for all sorts of economic studies at the regional level because they arrange important economic data in a logical pattern that can be used for impact studies, regional forecasts, decomposition analyses and much more. Unfortunately, the compilation of such tables from scratch is very expensive, because it entails a survey of a representative and sufficiently large sample of the regional industries. In order to save time and money, the usual approach nowadays consists of a *regionalization* of the national input-output table, which is available from statistical offices in most countries.

The regionalization can be performed by a number of methods. Until the 1990s, the most popular were the pure *nonsurvey methods*, as their requirements in terms of time, money and data are very low. However, they came under attack because the tables based on pure nonsurvey methods tend to be relatively inaccurate and, worse, systematically biased. In order to overcome the problems of pure nonsurvey methods without reverting to the expensive pure survey approach, researchers began to apply a mixture of the two, a so-called *hybrid approach*. This approach is now highly valued because it offers good empirical results at a reasonable cost.

Nevertheless, the importance of nonsurvey methods has not diminished, because a nonsurvey table may still permit useful insights despite its imperfections. Furthermore, the very popular 'top-down' hybrid approach is ultimately based on the application of a nonsurvey method. Therefore, the quest for improvements in the art of nonsurvey methods must continue, if only to improve the accuracy of the hybrid approach [Lahr, 1993].

With respect to the established nonsurvey methods, a rather sceptical view has taken hold. It has been shown that the empirical performance of nonsurvey methods is disappointing [Tohmo, 2004]. This is a consequence of their failure to acknowledge the role of cross-hauling. Moreover, their theoretical foundation is shaky as they are based on highly restrictive assumptions [Richardson, 1985]. Jensen [1990] therefore calls for an "extended debate of the theoretical and logical issues", which we attempt to contribute to with this paper. Departing from a review of the theoretical basis of input-output economics, we motivate the impression that nonsurvey methods have so far failed to incorporate cross-hauling because it manifests a violation of the homogeneous output assumption, a cornerstone of input-output analysis. Furthermore, as we argue in this paper, nonsurvey methods do not conform to ESA 95 accounting procedures, and they are incompatible with IO tables based on indirect allocation of imports, which limits their applicability to tables with direct allocation.

Based on these observations, we develop a new approach to the regionalization problem, the CHARM method. It is a pure nonsurvey method and closely related to the classical commodity balance approach. Compared to the established methods, it requires fewer restrictive assumptions, has a better theoretical foundation, and is compatible with ESA 95. Most notably, it incorporates the cross-hauling of commodities, based on an estimate of product heterogeneity, which should favourably affect its empirical performance. Like other nonsurvey methods, it may serve as a basis for the top-down hybrid approach.

In the next section, we revisit the layout and the logical structure of the input-output table. This may seem like a tedious exercise for some readers, but it is useful to clarify some basic issues which become important in our discussion of technical coefficients, ESA 95 rules, and the allocation of imports. In section 3, we review the commonly used nonsurvey methods and discuss how they contribute to the top-down hybrid approach. In section 4, we motivate the CHARM method and present its application. Section 5 concludes.

II The Structure of an Input-Output Table

In this section we present the setup of an input-output table and discuss some variations in the underlying accounting conventions, such as the treatment of imports. Readers with experience in input-output analysis are likely to know most of what is said in this section. Nevertheless, we consider it worthwhile to revisit these basic issues, because they will be important in the later discussion. A review of the basic issues also serves to introduce the mathematical notation used in this paper.

II.1 Definitions and Identities

The input-output table as we know it today was developed by Wassily Leontief, who was inspired by François Quesnay's concept of the *Tableau Economique*. This concept is based on the observation that an economy consists of a number of different sectors, each of which requires resources from the other sectors to produce some output. Economic transactions occur whenever the output from some sector is used as an input in another (or the very same) sector. These transactions between (or within) sectors may be observed and written down in an interindustry transactions matrix.

		Interindustry deliveries to sectors			
		1	2	…	n
Interindustry deliveries from sectors	1	$z_{1,1}$	$z_{1,2}$	…	$z_{1,n}$
	2	$z_{2,1}$	$z_{2,2}$	…	$z_{2,n}$
	…	…	…		…
	n	$z_{n,1}$	$z_{n,2}$	…	$z_{n,n}$

Table 1: The interindustry transactions matrix

Mathematically, the interindustry transactions matrix is denoted by *Z*. Its structure is shown in Table 1. Each element in the interindustry transactions matrix tells us the amount of sector *i*'s output which was delivered to sector *j* to be used as an input for that sector's production. For reasons which will become clear soon, it is useful to define an input-output coefficient $a_{i,j}$, which expresses the amount of input *i* used in the production of output *j* as a percentage of the total supply S_j of output *j*:

$$(1) \qquad a_{i,j} = \frac{z_{i,j}}{S_j}$$

Let Z_i^S denote the total amount of interindustry *sales* of sector *i*. This is found by adding over all cells in the *i*'th row. Using equation (1), this can be written as:

$$(2) \qquad Z_i^S = \sum_{j=1}^{n} z_{i,j} = \sum_{j=1}^{n} a_{i,j} \cdot S_j$$

In matrix notation, this can be conveniently expressed as:

$$(2.A) \quad Z^S = AS$$

Conversely, Z_j^P, the amount of interindustry *purchases* by sector *j*, is found by adding over all cells in the *j*'th column:

(3) $$Z_j^P = \sum_{i=1}^{n} z_{i,j} = \sum_{i=1}^{n} a_{i,j} \cdot S_i$$

In matrix notation, denoting the transpose of *A* by A', this becomes:

(3.A) $$Z^P = A'S$$

The reader may notice that in most introductory textbooks, Z^S is defined as the product of the matrix *A* and the vector of sectoral production *X*. This is indeed the case if imports are entered as (negative) components of final demand as suggested by Leontief [1986, p. 27]. In this paper, however, we follow the definitions of the ESA 95 standard, in which imports are entered as inputs rather than components of final demand [Eurostat, 1996]. Accordingly, the total supply of each output is equal to the sum of domestic production *X* and imports *M*:

(4) $$S = X + M$$

This explains why *S* rather than *X* appears in equations (3.A) and (3.B).

Domestic production of each output is equal to the sum of all intermediate inputs and value added *W*:

(5) $$X = Z^P + W$$

Using equations (3), (4), and (5), we can write total supply of output *j* as:

(6) $$S_j = \left(\sum_{i=1}^{n} a_{i,j} S_j \right) + W_j + M_j$$

This can be expressed more conveniently in matrix notation:

(7) $S = A'S + W + M$

We call this the *supply identity*, which states that the total supply of each commodity is equal to the sum of domestic production, which is in turn the sum of intermediate inputs and value added, and imports.

The amount supplied by sector *j* can be used either as intermediate inputs in other sector or as final use, for example for consumption or investment. Thus, total use U_i of sector *i*'s output is equal to the sum of intermediate use in all other sectors and final use Y_i:

(8) $U_i = Y_i + \sum_{j=1}^{n} z_{i,j}$

It is customary to divide final use into domestic final use D_i and use for exports E_i:

(9) $U_i = D_i + E_i + \sum_{j=1}^{n} z_{i,j}$

Note that the terms in the sum sign in equation (10) are simply the row total of matrix *Z* or *AS*. Thus, the equation can be written in matrix terms as:

(10) $U = D + E + AS$

This is what we call the *demand identity*. It states that total use is equal to final use plus exports plus intermediate use.

The beauty of an input-output table lies in its ability to express the supply and demand identities in the form of a well-structured table. Table 2 provides an example of such an IO-table, assuming that the economy consists of only two sectors.

		Intermediate use…		Final use		Total use
		in sector 1	in sector 2	Domestic	Exports	
Intermediate supply…	from sector 1	$a_{1,1}S_1$	$a_{1,2}S_2$	D_1	E_1	U_1
	from sector 2	$a_{2,1}S_1$	$a_{2,2}S_2$	D_2	E_2	U_2
Other supply	primary inputs (value added)	W_1	W_2			
	Imports	M_1	M_2			
Total supply		S_1	S_2			

Table 2: Input-output table for a two-sector economy

Each row of the IO-table contains the demand identity for a single sector, and each column contains the supply identity for a single sector. For example, row 1 may be read as:

(11) $$a_{1,1}S_1 + a_{1,2}S_2 + D_1 + E_1 = U_1$$

This is exactly the demand identity as expressed in equation (9). In a similar vein, column 1 may be read as:

(12) $$a_{1,1}S_1 + a_{2,1}S_1 + W_1 + M_1 = S_1$$

This is exactly the supply identity as expressed in equation (6). Finally, we may note that total supply must be equal to total use (which includes changes in stocks):

(13) $$S = U$$

This implies that the IO-table presented above is *symmetric* in the sense that each element in the total supply row is equal to the corresponding element in the total use column.

II.2 Impact Analysis and Multipliers

While even a casual glance at the input-output table itself may offer interesting insights, its most common use lies in impact analysis by means of multipliers. The aim of impact analysis lies in assessing the effect of an exogenous change in final use (the eponymous 'impact'). Classical examples of this include government spending on infrastructure projects or a change in the consumption expenditures of households. In terms of the input-output table, such exogenous impacts are reflected in changes in the vector of final use, *D*. Under a certain set of assumptions, we may use the data contained in the input-output table to assess the impact of that change on each sector's output.

The *homogeneous output assumption* states that each sector produces only one kind of output. Thus, if there is a sector called "ice cream production", there can be only one kind of ice cream, and no distinction may exist between chocolate, vanilla, or strawberry ice cream. Next, the *unique origin assumption* states that each output commodity is produced in only one sector. That is, ice cream may be produced only in one sector, so all firms or companies which produce ice cream should be allocated to the same sector. Furthermore, the *linear technology assumption* states that in each sector output is produced using a linear production function. By this we mean that each sector's output is related in fixed proportions to each individual production input and vice versa, which implies that the input-output coefficients $a_{i,j}$ are fixed. Such a production technology is sometimes called a Leontief production function.

These assumptions are of course not met in reality, where issues such as product differentiation, joint production, economies of scale and input substitution complicate the picture. Nevertheless, insofar as they allow a reasonable *approximation* of developments in the real world, they form a valid basis for economic theorizing and empirical analyses. Some violations of these three assumptions are discussed below.

If we accept the aforementioned assumptions, we can use the data from the input-output table to analyse the impact of an exogenous change in final demand in great detail. For example, if consumers decide to buy one additional unit of commodity *i*, we know from the unique output assumption that this commodity is produced in only one sector, namely sector *i*. Thus, the *direct effect* of the demand impulse is that the supply of sector *i* must increase by one unit to satisfy the increased demand.

However, the increased supply of sector *i* can only be achieved if sector *i* receives more inputs. The homogeneous output assumption together with the linear technology assumption implies that if sector *i*'s output increases by one unit, each input *j* must increase by an amount $a_{i,j}$. These inputs, however, have to come from somewhere. Other sectors have to increase their supply as well, which requires for them to use additional inputs again. Through these *indirect effects*, the direct effect is multiplied, so the total supply in the economy must increase by more than one unit.

The size of this multiplier effect can be calculated using relatively simple matrix algebra. Note that since $U = Z^P + Y = AS + Y$, equation (13) can be written as $S = AS + Y$. This equation can be solved for *S* as a function of *Y*:

$$\text{(14)} \qquad S = (I - A)^{-1} Y$$

I denotes the identity matrix. Equation (14) is probably well-known to many readers with *X* in place of *S*, which is again due to the fact that most input-output textbooks define imports as negative components of final demand, whereas ESA 95 rules define them as inputs.

Assuming that the proportion of imports *M* to domestic production *X* remains constant, any change in the supply vector *S* will lead to proportional changes in *X* and *M*. Thus, the change of each sector's domestic production can also be estimated, and from this information inferences may be drawn on the changes in the components of value added, for instance the income of workers. Using further statistical data, the effects on employment, environmental pollution etc. can be estimated. Thus, impact analysis using input-output techniques is routinely applied to a very wide range of economic issues. The results of such an impact analysis depend to a large extent on the estimated values of the multipliers. Therefore, it is important to achieve a high accuracy when estimating these multipliers.

II.3 Sector Classification by Industry or Product

Under the assumptions mentioned above, each industry produces exactly one type of product, and each product in turn is produced by only one single industry. In this case, each 'sector' in the input-output table constitutes one industry producing one type of product. Because of this, the terms 'sector' and 'industry' are quite frequently used as synonyms in the input-output literature.

Reality, however, is much more complex. Many industries produce more than one product, and most products are produced by more than one industry. Therefore, great care is required in defining the boundaries of a 'sector' and allocating the observed production processes and transactions to these sectors. In the past, different approaches have been used. Sectors may be classified by industry or by commodity (i.e. product). In the former approach, the output of each firm is allocated to the industry to which that firm belongs. In the latter approach, each firm's output is split up between the different types of products and allocated to the sector defined by the type of product. A concise exposition of these allocation procedures is provided in chapter 5 of the book by Miller and Blair [1985].

According to ESA 95, the symmetric input-output table is arranged on a by product basis. The allocation of each industry's output is performed by means of a *supply table*. This table allocates the output of each industry to one or more product categories. Thus, each sector is defined as a product category.

II.4 Allocation of Imports

As mentioned above, the early input-output literature usually treated imports as a component of final use. In the input-output tables, they were entered as a column of negative entries. Currently, however, the more widespread practice appears to be the one suggested by ESA 95. According to that practice, imports are treated as inputs and accordingly entered as a row in the input-output table.

Nevertheless, there are still two ways of defining imports. Under *direct allocation* of imports, any imported products are allocated as inputs to the sectors which actually purchase them. This means that imports appear only at the bottom of the input-output table, below primary inputs, and not in the interindustry transactions matrix. Under *indirect allocation*, each imported product is entered as an input to the sector which would have produced it if production had occurred within the country. Furthermore, it is also entered in the interindustry transactions matrix as a transaction from the sector which would have produced the product domestically to the sector which uses it. In the parlance of ESA 95, the indirect allocation table seems to be the standard input-output table, whereas the direct allocation table is referred to as the 'input-output table for domestic output' [Eurostat, 1996].

It must be pointed out that there is no 'preferred' way of allocating imports. Both direct and indirect allocation approaches have their merits; for some research questions a table based on direct allocations is more useful while for other questions the indirect allocation table is more useful. Although Jensen [1978] shows that the indirect allocation multipliers tend to be larger than the direct allocation multipliers, this does not mean that either of them are 'wrong'. However, they have to be interpreted appropriately: direct allocation multipliers show the effect on domestic *output*, whereas indirect allocation multipliers show the effect on *supply*, which may originate from both domestic production and imports. Therefore, it may be appropriate to introduce the term *supply multipliers* for the latter in order to avoid confusion.

Consider the process of calculating the direct coefficients matrix *A*. The matrix *A* consists of the input-output coefficients $a_{i,j}$, which are defined in equation (1) as $a_{i,j} = z_{i,j} / S_j$. However, the concept of interindustry transactions, $z_{i,j}$, has a different interpretation depending on the allocation of imports. If sector *j* imports some amount of commodity *i*, a direct allocation table records this only as an import of sector *j*. Mathematically, it is included in M_j but not in $z_{i,j}$. Under indirect allocation, by contrast, such imports are recorded as part of $z_{i,j}$ and also as an import to sector *i*, so they are added to M_i.

The input-output coefficients $a_{i,j}$ are sometimes called *technical coefficients*. Strictly speaking, this is true only if imports are treated as components of final demand, which is not the case under ESA 95 rules. Therefore, the $a_{i,j}$ coefficients that are calculated from ESA 95 tables cannot be interpreted as technical coefficients. Note that it is not possible to calculate technical coefficients from direct allocation tables, because those tables do not contain the necessary information – imported products are lumped together in an amount of 'imports', and it is impossible to tell whether these imports consist of crude oil, vegetables, or services. However, it is possible to calculate technical coefficients from indirect allocation tables by means of the following formula:

$$(15) \qquad a_{i,j}^X = \frac{z_{i,j}}{X_j} \qquad \text{or} \qquad A^X = ZX^{-1}$$

The superscript *X* denotes the fact that the technical coefficient $a_{i,j}^X$ is defined as the ratio of the interindustry transaction $z_{i,j}$ to the domestic production, X_j, rather than supply, S_j.

II.5 Aggregation Level

An important question in the construction of input-output tables is the level of sectoral disaggregation which should be used. If the level of disaggregation is high, the number of sectors is large, and vice versa. As a rule, higher disaggregation is desirable because if the economy is divided into many small sectors, the three fundamental assumptions are more likely to hold. For example, the output of a sector 'vehicle production' is probably very heterogeneous, but if such a sector is disaggregated into several smaller sectors such as 'automobile production', 'truck production', 'motorcycle production' and so on, the output if those smaller sectors resembles more closely the homogeneous output ideal.

Historically, there have been two obstacles against the construction and application of highly disaggregated input-output tables: high computational requirements and data acquisition costs. The first obstacle has been considerably reduced by the widespread diffusion of powerful personal computers, but the second obstacle remains. A high disaggregation level requires more data to be collected, which can be prohibitively expensive.

Furthermore, input-output analysis is frequently applied to regional issues, and according to Lahr [1993] some authors seem to have argued that “smaller models are the best for small regions”. However, it has been shown that aggregation, in addition to obscuring sectoral details, induces an aggregation bias, which leads Lahr [1993] to the conclusion that even in regional analyses the input-output table should remain “as detailed as possible” and, hence, at a high level of disaggregation.

II.6 Applications

A purely subjective glance at the list of input-output publications in recent years gave us the impression that input-output analyses are nowadays mostly conducted in two research areas: regional economics [Augurzky & Neumann, 2005, Midmore et al., 2006], environmental and resource economics [Llop, 2007], or a combination of both [Dietzenbacher & Velázquez, 2007]. The reason for this appears to be the great sectoral detail offered by input-output tables. In environmental and resource economics, one is often interested in the effects of consumption changes on environmental pollution or resource consumption. Using input-output techniques, such changes can be traced throughout the economy and decomposed into direct and indirect effects. In regional economics, one usually needs a very detailed analysis of a small regional economy, and such detail may be offered by the multi-sectoral nature of input-output analysis. Furthermore, the three fundamental assumptions mentioned above may be more realistic for a small regional economy. For example, although the agricultural sector of the US is not at all homogeneous due to different climate zones, the agricultural sector of Indiana is much more homogeneous, consisting mostly of corn and soybeans.

However, the application of input-output techniques to regional economic issues is often hindered by the fact that input-output tables are only available for national, not regional economies. In such cases, the national table must be adjusted to reflect the specific aspects of the region in question. The methods for this regionalization are discussed in the following section.

III Regionalization Methods

In principle, a regional input-output table (henceforth RIOT) can be constructed in two ways: Either one can collect all the necessary data by means of a survey, or one can use statistical information which has already been published by other sources. The former approach is called the *survey method*, and the latter the *nonsurvey method*. Since conducting a survey with the necessary detail is extremely expensive, the pure survey method is rarely used nowadays. Instead, *hybrid methods* have become popular which are based on some mixture between the pure survey and pure nonsurvey approach. In this section we will first discuss nonsurvey methods, then hybrid methods.

III.1 Nonsurvey Methods

The most widely used nonsurvey methods are based on the location quotient (LQ) approach, which is why that approach receives the greatest attention in this section. Other methods include the commodity balance (CB) or supply-demand pool approach, iterative methods (RAS), and econometric estimation. We will first discuss location quotient methods.

III.1.1 Location Quotient Methods

The location quotient method was originally developed as a tool for economic base analysis, where it was used to indicate whether a certain industry was export- or import-oriented. This section presents the simple location quotient method in some detail because it is the most widely applied version. A number of different variations have been developed to overcome the theoretical shortcomings of the simple method, but empirical evidence suggests that these do not necessarily perform any better in terms of accuracy. Therefore, these variations are only sketched rather briefly.

All location quotient methods are based on the assumption that any regional input-output coefficient $a_{i,j}^{R}$ can be related to its national counterparts $a_{i,j}^{N}$ in the following fashion:

$$(16) \qquad a_{i,j}^{R} = t_{i,j} \cdot a_{i,j}^{N}$$

Thus, the regional IO-coefficient is proportional to its national counterpart. The factor of proportion, $t_{i,j}$, is sometimes interpreted as a 'trading coefficient' [Round, 1983] or a 'regional purchase coefficient' [Stevens et al., 1983].

At this point it is crucial to recall our discussion on the interpretation of $a_{i,j}$. If imports were recorded as components of final demand, it would be correct to interpret $a_{i,j}$ as a technical coefficient. According to ESA 95 rules, however, imports are recorded as inputs, so $a_{i,j}$ cannot be regarded as a technical coefficient. Under direct allocation, $a_{i,j}$ does not represent the entire amount of input *i* used in the production of output *j*, because some of input *i* may be obscured in the import row. Under indirect allocation, $a_{i,j}$ includes the entire amount of input *i* used in the supply of output *j*, but supply consists of both domestic production and imports.

Anyhow, the literature seems to suggest that if the regional economic structure is identical to the national one (the regional economy is a scaled-down version of the national economy), all the trading coefficients will be equal to one, and $a_{i,j}^{R} = a_{i,j}^{N}$ for all *i* and *j*. In this case, the regional direct coefficient matrix is equal to its national counterpart. Therefore, all differences between the regional and national matrices must come from differences in the trading coefficients.

According to the simple location quotient method, $t_{i,j}$ is a function of the relative size of industry *i*. Mathematically, the simple location quotient (SLQ) is defined as:

$$(17) \qquad SLQ_i = \frac{L_i^R / L^R}{L_i^N / L^N}$$

L^R is total employment in the region, L^N is total employment in the nation, and the subscript *i* refers to a certain industry *i*. Ideally, one would use output rather than employment, but since employment data is usually available from statistical offices whereas output data is not, most applications use employment as a proxy for output, which requires the assumption that labour productivity in the region is the same as in the nation as a whole. Equation (17) shows that SLQ is calculated by dividing industry *i*'s share in total employment at the regional level by its share in total employment at the national level.

The adjustment procedure as outlined by Schaffer and Chu [1969] proceeds as follows: If SLQ_i is smaller than one, "local production is assumed to be inadequate to supply local needs – no exports can be made and imports are necessary". SLQ_i is substituted for $t_{i,j}$ in equation (16), so $a_{i,j}^{R}$ will be smaller than $a_{i,j}^{N}$. On the other hand, "a location quotient equal to one means that the region is self-sufficient", and "a location quotient greater than one means that the region exports some of output *i*" [Schaffer & Chu, 1969]. In those cases, $t_{i,j}$ is assumed to be equal to one, and $a_{i,j}^{R} = a_{i,j}^{N}$.

It is important to realize that this line of reasoning is valid only if imports are allocated directly. Under indirect allocation, there is no reason to expect that the size of industry *i* should have any effect on $a_{i,j}^{R}$. If industry *i* is too small to fulfil industry *j*'s demands, industry *j* will certainly import some amount of output *i*, but under indirect allocation this will not be reflected in $a_{i,j}^{R}$ but rather in M_i. Therefore, location quotient methods are applicable only to direct allocation tables.

Of course, one may expect that if industry *i* is too small to fulfil industry *j*'s demand, industry *j* might substitute other inputs for input *i*, so $a^R_{i,j}$ would be smaller even under indirect allocation. This, however, would mean that industry *j* is using a different technology. While this is perfectly plausible according to common sense, it does not help the SLQ method at all, because location quotient methods are based on the assumption of identical technology anyway. Thus, the SLQ method, as well as all other variants of the location quotient method, is only applicable to direct allocation tables.

Even when applied to direct allocation tables, the SLQ method is valid only under very restrictive assumptions. For example, it is implicitly assumed that regional demand for each commodity is proportional to the region's size: If the regional economy represents 10 percent of the national economy, then demand for industry *i*'s output is assumed to be 10 percent of national demand. If this is true, then the regional industry needs to produce 10 percent of the national total in order to be self-sufficient, which is the logic underlying the SLQ approach. However, it is possible that regional demand for commodity *i* is less (or more) than 10 percent of the national total. This can happen, for instance, if the structure of regional final demand is different, or if the industries which use industry *i*'s output are underrepresented (or overrepresented) in the region.

In order to control for the latter possibility, the purchase-only location quotient (PLQ) was developed [Consad Research Corporation, 1967]. Its calculation is similar to that of SLQ, but instead of regional and national total employment, only employment in industries which use the output of industry *i* is considered. A different approach to the same problem is embodied in the cross-industry location quotient (CILQ) approach. The inventor of this approach appears to have been Charles Leven [Schaffer & Chu, 1969]; a concise exposition (in Dutch) is provided by Klaassen and Verster [1974]. The CILQ is calculated as follows:

$$(18) \qquad CILQ_{i,j} = \frac{L_i^R / L_i^N}{L_j^R / L_j^N}$$

The reasoning behind this approach is that if industry *j* is underrepresented in the region, a small industry *i* will be enough to fulfil its demands and may actually be export-oriented despite its being underrepresented as well, as long as it is less underrepresented than industry *j*. The CILQ approach requires more computations, because whereas the SLQ and PLQ approaches calculate only one sufficiency indicator for each industry, CILQ calculates one for each cell of the matrix *A*.

In order to test the performance of these different LQ methods, Morrison and Smith [1974] applied them to estimate a regional IO-model for a small city in England and compared the results to an empirically derived model for the same city. They concluded that “the most simple of the tested methods (SLQ) emerges, on the whole, as the best of the purely nonsurvey approaches”. More recent evidence [Bonfiglio, 2005] seems to suggest that PLQ may perform slightly better than SLQ, but the difference is not great. CILQ, despite being more convincing than SLQ from a theoretical perspective, appears to be clearly inferior in terms of empirical results. In the study by Morrison and Smith the performance of the basic CILQ method was disastrous; they proposed a modified CILQ method whose performance was better but still unsatisfactory.

A common trait of all the LQ methods presented so far is that they tend to underestimate the volume of interregional trade. This is quite unfortunate because it implies a bias in the multipliers that are calculated from IO-tables based on those methods [Tohmo, 2004]. In general, one may expect that smaller regions trade relatively more than large regions. Therefore, Round [1978] proposes a new formula for the estimation of trading coefficients which takes the size of the region into account. This method is known under the label ‘semi-logarithmic location quotient method’, or simply RLQ. Building upon Round’s work, Flegg et al. [1995] propose the so-called FLQ formula to estimate trading coefficients. This formula in a way ‘corrects’ for region size by means of a parameter. Unfortunately, the value of this parameter cannot be known beforehand and requires, in fact, an educated guess.

Bonfiglio [2005] tests these methods in an exercise similar to that of Morrison and Smith, using data for the region of Marche in Central Italy. He finds that PLQ yields the best performance in terms of reproducing a survey-based IO-coefficients matrix, followed by SLQ and, with some distance, FLQ. In terms of reproducing survey-based multipliers, FLQ may outperform SLQ and even PLQ if the parameter value is chosen correctly, but if it is chosen incorrectly FLQ performs worse than SLQ and PLQ. In nearly all cases, the performance of FLQ is significantly better than either RLQ or CILQ.

Tohmo [2004] performs similar tests on the Finnish region of Keski-Pohjanmaa and concludes that SLQ, CILQ and RLQ “tend to produce substantially overstated regional multipliers”, whereas the FLQ approach yields “more appropriate estimates”. Riddington et al. [2006], by contrast, apply LQ methods to Scottish regional data and find that the performance of FLQ is inferior to that of SLQ and CILQ. Thus, the general picture is not quite clear, and none of the LQ methods seems to be clearly superior to the others.

All of them, however, have been subject to harsh criticism questioning the validity of the LQ approach in general. According to Richardson [1985], LQ methods require four critical assumptions to hold:

1) Equal labour productivity in the region and the nation
2) Equal consumption per employee in the region and the nation
3) If the region exports a commodity, it does not import the same commodity
4) The nation neither imports nor exports any given commodity in net terms

Because these assumptions are very unrealistic, argues Richardson, the LQ approach tends to produce very poor results. In particular, "the technique grossly underestimates exports and, hence, overestimates multipliers". Furthermore, he claims that the LQ method has "no sound theoretical grounds", and "the results obtained with it are unconvincing" [Richardson, 1985].

Further criticism against traditional LQ methods is expressed by Maks and Oude Wansink [1998]. They point out that even if a location quotient is calculated to be larger than one, the trading coefficient is not allowed to be larger than one. From theory, however, there is no reason that would preclude a regional trading coefficient from being larger than one. Therefore, Maks and Oude Wansink propose a variant of the CILQ method which allows the trading coefficient to be between zero and two. This is certainly an improvement, but the value of two is still an arbitrary cut-off point. The problem is thus meliorated but not completely solved.

III.1.2 Commodity balance or supply-demand pool method

An alternative regionalization method is provided by the commodity balance (CB) or supply-demand pool (SDP) approach, based on the work by Isard [1953]. It presumes that regional production and regional consumption of each commodity can be estimated. The difference between the two is called the 'commodity balance', although the term 'net exports' may be more popular in contemporary parlance.

In some cases, data on the regional production X_j^R of an industry *j* may be available. If not, it may be estimated from regional employment data, which is more often available, assuming that labour productivity in the region is equal to the national average. Under the assumption of identical technology, the input requirements of the regional industries can be estimated using national technical coefficients and regional output estimates:

$$(19) \qquad z_{i,j}^R = a_{i,j}^X X_j^R$$

Note that when we are talking about technical coefficients, we must use the coefficients defined by equation (16). Using (19), the regional interindustry transactions matrix Z^R can be estimated. The row totals of that matrix constitute the regional intermediate use of the individual commodities, Z^{PR}.

Next, final use (excluding exports) of each commodity in the region is estimated by assuming proportionality with its national counterpart. Thus, we have an estimate for D^R. Using the facts that, by definition, $S^R = X^R + M^R$ and $U^R = Z^{PR} + D^R + E^R$ along with the market clearing condition (13), we can write:

$$(20) \quad B^R \equiv E^R - M^R = X^R - (Z^R + D^R)$$

B^R, the 'commodity balance', is defined as the difference between regional exports and regional imports. Equation (20) states that this must be equal to regional production minus regional consumption, which is the sum of regional intermediate use and regional final use.

One problem of the CB approach, however, lies in its inability to estimate actual exports and imports – it can only estimate *net* exports. The construction of a full regional input-output table, however, requires information on both exports and imports. Moore and Petersen [1955] solve this problem by assuming that if B^R is positive, M^R equals zero and E^R equals B^R. If B^R is negative, they assume that E^R equals zero and M^R equals B^R [Schaffer & Chu, 1969]. In effect, as Round [1983] puts it, "net export, or import, is assumed to equal the net surplus, or deficit, as the case may be".

Although the philosophy underlying the CB and LQ approaches appears very different at first sight, they are formally very closely related and tend to produce similar results [Round, 1983]. Like the LQ approach, the CB approach tends to underestimate regional trade and thereby leads to overestimated regional multipliers. Among the different LQ methods, the CILQ resembles the CB approach most closely [Round, 1972]. Since that is one of the rather poorly performing LQ methods, it comes as no surprise that the CB approach performs considerably worse than most LQ approaches, at least for single-region models [Morrison & Smith, 1974, Bonfiglio, 2005]. However, it does have an advantage in multi-region models because of built-in consistency checks because the exports and imports of the individual regions within a country have to balance each other [Richardson, 1985].

III.1.3 Econometric estimation of regional purchase coefficients

The biggest drawback of the approaches discussed so far is that they calculate interregional trade as a residual after regional production and consumption as well as intraregional transactions have been estimated. Naturally, this results in rather poor estimates of interregional trade, which, as we have seen, leads to biased estimates of multipliers and raises questions about the usefulness of the entire nonsurvey exercise.

A very promising approach to achieving better estimates of interregional trade is presented by Stevens et al. [1983]. They argue that the 'regional purchase coefficient' (RPC) is a function of the relative unit cost in different regions and shipment cost, which in turn depends on the distance between two regions and the weight/value ratio of the commodity in question. Using a unique set of interregional transportation data, Stevens et al. then compute RPC estimates by means of regression analysis. Those RPC estimates are then used to estimate the regional trade patterns.

Although the approach by Stevens et al. appears to be more convincing than LQ and CB approaches from both a theoretical and empirical perspective, it has never found widespread application. The reason is obviously that it requires a large amount of data. If such data happens to be available, the RPC approach is certainly very attractive. Unfortunately, however, this is no longer the case, because one of the most important data sources, the Census of Transportation, has been discontinued [Richardson, 1985]. As a result, the RPC approach seems to have fallen into oblivion. The regionalization method by Gabriel [2001] seems to have drawn inspiration from the RPC approach, but because of the non-availability of data, it boils down to making informed guesses rather than estimates of the value of 'regional preference factors'.

III.1.4 Iterative Method (RAS)

The RAS method was originally developed as a tool for updating aging input-output tables. However, it has also found widespread application as a regionalization tool. Its application is thoroughly explained, for instance, by Miller and Blair [1985, pp. 276ff], so it will not be repeated here. Compared to the LQ and CB approaches, RAS has the disadvantage of requiring more data, because the row and column totals (Z^{RP} and Z^{RS}) of the interindustry transactions matrix are presumed to be known.

Furthermore, while the theoretical justification of the LQ and CB approaches may rest on questionable assumptions, RAS as a regionalization method possesses no such theoretical justification at all. It is a purely mechanical adjustment process. If applied to temporal updating, it may be argued that the adjustments to the coefficients reflect substitution and fabrication effects [Stone, 1961], but if applied to regionalization, no such interpretation seems feasible [Richardson, 1985].

Assuming that data on Z^{RP} and Z^{RS} have somehow been acquired, the RAS method makes up for these disadvantages by providing more accurate estimates than the LQ and CB methods [Morrison & Smith, 1974]. However, since RAS requires (and uses) much more information than the other methods, there is no basis for a 'fair' comparison. Lahr [1993] even argues that since the data required for RAS is usually not available, and would have to be collected by means of a survey, RAS should not be considered a nonsurvey method. In his view, RAS may constitute a complement to nonsurvey methods rather than a substitute. He proposes its application as a tool for balancing a table as the final step of a so-called *hybrid approach*. This approach is discussed in the following subsection.

III.2 Hybrid Methods

Facing a tradeoff between the excessive cost of the pure survey approach on the one hand and the limited accuracy of the pure nonsurvey approach, input-output analysts were attracted by a mixture between the two, which has become known as the *hybrid* approach. Many different varieties of this approach have been developed; Lahr [1993] provides an extensive survey and a more recent update [Lahr, 2001b]. Using the terminology of West [1990], hybrid methods can be divided into 'bottom-up' methods, which use only information from the respective region, and 'top-down' methods, which use a national IO table as a point of departure. An example of the former approach is the DEBRIOT procedure outlined by Boomsma and Oosterhaven [1992]; an example of the latter is the GRIT approach, which is extensively used by a number of Australian input-output modellers. Since the focus of this paper is on non-survey methods, which are applied in the top-down approach but not in the bottom-up approach, the further discussion concentrates on the top-down approach.

In general, the top-down hybrid approach consists of five phases:

1) Use some nonsurvey method to regionalize the national IO table
2) Identify key sectors
3) Collect superior data for key sectors
4) Enter superior data into estimated table
5) Balance table

In the first phase, one of the traditional nonsurvey methods or a variation thereof is applied to the national IO tables. The GRIT approach, for example, uses a modified location quotient method [West, 1980]. The result of this exercise constitutes a 'first guess' of the regional IO table of limited accuracy and reliability. However, this 'first guess' is not interpreted as an actual approximation to the regional IO table. Rather, it is used in the second phase to identify 'key sectors' in the region.

The philosophy of the hybrid approach is to strike a balance between the accuracy of the table and the cost of constructing it. Jensen [1980] introduces the concepts of *partitive accuracy*, which refers to individual cells, and *holistic accuracy*, which refers to the overall accuracy of the input-output table as a whole. Naturally, if we are interested in the partitive accuracy of a few cells, superior data should be collected for precisely those cells. For the holistic accuracy of an input-output table, however, some cells are more important than others. Therefore, if holistic accuracy is the goal, and money and time are scarce resources, "it is inefficient to spread these resources evenly over all the cells in order to obtain superior or updated estimates; primary attention should be given to the key sections of the table" [West, 1981].

A key 'section' may be an individual cell or a group of cells. There are different procedures for the identification of them, based on measuring the forward and backward linkages of individual cells or entire sectors [West, 1982, Schintke & Stäglin, 1988, Lahr, 2001b]. The accumulated experience of RIOT construction has shown that superior data collection tends to be most fruitful in the household/labour sector, resource-using sectors such as agriculture and mining, and 'miscellaneous' sectors, which tend to be very heterogeneous. Therefore, Lahr [1993] suggests that „hybrid model constructors should pursue the most accurate non-survey model of their region as possible (use accurate regional purchase coefficients and minimize aggregation), always seek superior data for households and establishments in resource-based and 'miscellaneous' sectors, and sequentially identify other sectors that should receive superior data."

In the third phase, data is collected for individual cells or sectors. Under efficiency considerations it makes sense to identify key sectors rather than cells in phase 2, because it is relatively cheaper to ask one firm about all its inputs, which yields a column estimate, than to ask firms in different sectors about a small number of different inputs, which yields estimates for individual cells. Therefore, "identifying sectors rather than individual cells as targets of survey work is likely to prove more cost-effective" [Lahr, 2001b]. The fourth phase consists of entering the data thus acquired into the table which was estimated in the first phase.

In phase 5, the data in the estimated table must be reconciled and balanced. Reconciliation of data may be necessary if the survey yields contradictory results. For example, firms may have been asked how much they bought from other sectors and how much they sold to those other sectors. The former question yields data by column, the latter by row. If more than one sector is surveyed, the analyst will acquire information on some cell entry $z_{i,j}$ from two different sources: An estimate on what sector *i* sold to sector *j*, and an estimate on what sector *j* bought from sector *i*. Generally, the two estimates will not coincide, and some form of reconciliation is called for. The reconciliation of these two estimates usually requires an informed judgment by an expert, although statistical methods have also been proposed [Round, 1983].

Finally, the table needs to be balanced to ensure that column totals equal row totals. This can be achieved by calculating some component of final demand as a residual. Alternatively, the RAS procedure may be applied [West, 1980]. If there are still resources (time and money) available, the procedure from phase two onward is reiterated until the analyst decides that the marginal cost of another iteration exceeds the marginal benefit or runs out of resources.

III.3 Scope and Usefulness of Regionalization Methods

As argued above, pure nonsurvey methods are characterised by a lack of theoretical backing and tend to produce rather disappointing empirical results. Nevertheless, this does not mean that they should be discarded completely. Even though nonsurvey tables may not be as accurate as we wish, they still perform “a reasonable job in identifying sectors that are most important to target for superior-data collection” [Lahr, 2001b]. Thus, nonsurvey methods are still a useful tool, if not alone then certainly as part of a hybrid approach.

Currently, the top-down hybrid approach appears to be the most popular method of constructing regional input-output tables due to its attractive combination of survey and nonsurvey elements. Nevertheless, there is still scope for improvement. West [1990] argues that “the single most important factor determining the final accuracy of the table is the availability and use of superior data” [West, 1990]. It appears that subsequent research has heeded this recommendation, and efforts have indeed been directed at optimizing the later phases of the hybrid procedure.

However, the importance of the nonsurvey method should not be underestimated. Lahr [1993] emphasizes that “since hybrid IO models are based on non-survey models, it is critical to use the best non-survey methods possible”. As argued above, however, it is not clear which of the nonsurvey methods is the ‘best’ one. Actually, all of them are quite disappointing because they fail to take into account two important features of regional economies. According to Lahr [2001b], “differences in *trade patterns* and *technology* account for most of the differences between the direct requirements matrices of a nation and each of its component regions” (emphasis added). Therefore, these two factors should receive more attention.

With respect to technology, it has been common practice to assume invariant technology across a nation since the early work by Isard [1951]. Unfortunately, this problem is hard to deal with, because estimating regional technical coefficients would require a large amount of data which, if acquired, would basically permit the construction of a survey table, and nonsurvey methods would be redundant. However, Lahr [2001a] proposes that one should use labour income rather than employment as an indicator for sector size. Although this procedure cannot account for differences in the ‘production recipe’, it can at least correct for overall productivity. Furthermore, empirical evidence seems to indicate that differences in trade patterns seem to be much more important than differences in technology. Harris and Liu [1998], referring to a study by Park et al. [1981], even argue that “the effect of errors in the technical coefficient matrix on the overall accuracy of the model is surprisingly negligible”.

Therefore, the most urgent problem seems to be the misrepresentation of trade patterns. A central role is played by cross-hauling or intraindustry trade. As pointed out by Lahr [2001a], traditional nonsurvey methods “cannot estimate both in- and outflows and estimate only *net* outflows. This is because such techniques cannot permit the crosshauling of commodities that pervades all existing data on interregional trade".

The non-incorporation of cross-hauling, a general feature of traditional nonsurvey methods, is carried over into the top-down hybrid approach. Therefore, what Richardson [1985] says about the nonsurvey approach is just as true for the hybrid approach: "The most critical need to salvage the location quotient and similar approaches is to develop improved nonsurvey adjustments that correct for the effects of cross-hauling". Since the overall performance of the hybrid approach hinges on the accuracy of the 'first guess' on which it is based, a nonsurvey method which allows for cross-hauling would most likely improve the usefulness of the hybrid approach as a whole. The next section presents an attempt to incorporate cross-hauling in a pure nonsurvey method.

IV Adjusting for Cross-Hauling

In this section, we present a new approach to nonsurvey regionalization, which goes by the name CHARM (Cross-Hauling Adjusted Regionalization Method). The approach is basically a variant of the CB approach, but it accounts for cross-hauling by estimating product heterogeneity. We first underline the importance of product heterogeneity for cross-hauling, the present the calculations required by CHARM. Finally, we show how CHARM may be implemented as part of a top-down hybrid approach.

IV.1 Why Does Cross-Hauling Occur?

Cross-hauling occurs mainly because one of the fundamental assumptions of input-output analysis, the homogenous output assumption, is violated in reality. If each product was completely homogeneous, as stated in the assumption, there would be no reason for the cross-hauling of commodities. For example, if automobiles were homogeneous, consumers in Lower Saxony, where the Volkswagen headquarters is situated, would buy only Volkswagen cars, and consumers in Bavaria, where DaimlerChrysler produces its Mercedes, would buy only Mercedes. In reality, however, automobiles are quite heterogeneous. Mercedes are shipped from Bavaria to Lower Saxony, and Volkswagens are shipped in the other direction. Cross-hauling occurs. Empirical evidence supports the view that product heterogeneity is the main reason for the cross-hauling of commodities. Harris and Liu [1998], for example, argue that "industries where product differentiation and brand preference are important usually exhibit considerable cross-hauling", citing a study by Norcliffe [1983] as a source for this observation.

The regional input-output literature has had great difficulty with cross-hauling because product heterogeneity constitutes a violation of one of the fundamental assumptions of input-output analysis, the homogeneous output assumption [Kronenberg, 2007]. In principle, the problem of cross-hauling "can be reduced by implementing the location quotient method using data at the most detailed level of industrial disaggregation possible" [Isserman, 1980], because at a very high level of disaggregation the product groups are relatively homogeneous. However, in most countries data are simply not available at a sufficient disaggregation level to preclude cross-hauling. In Germany, for example, the national input-output table contains no more than 71 sectors, among which are sectors such as "clothing" or "machines". Those sectors are certainly characterised by a lot of heterogeneity, and cross-hauling is definitely occurring. Therefore, cross-hauling can never be excluded, and must be taken into account.

Another reason for cross-hauling may be proximity to a border. If a firm is located close to a border, it may be the case that the closest supplier of a certain input happens to be situated on the other side of the border. This is likely to occur more frequently in small regions, where a larger share of the total area is close to the border. It is for this reason that Flegg et al. [1995] propose to adjust the estimated trading coefficient for region size. However, as argued above, their approach has met with mixed success, so it may be useful to concentrate on product heterogeneity rather than region size.

IV.2 Estimating Product Heterogeneity

Since cross-hauling is a function of product heterogeneity, it is in principle possible to estimate the degree of product heterogeneity if cross-hauling is observed. However, cross-hauling does not depend on product heterogeneity alone: If a region does not consume a certain product, it has no reason to import that product, no matter how heterogeneous that product may be. Also, if a region does not produce a certain product, it has no reason to engage in cross-hauling of that product; it is more likely to simply import that product until demand is fulfilled. For these reasons, we assume that cross-hauling is a function of product heterogeneity, domestic (or regional) production, and domestic (or regional) final and intermediate use:

(21) $$CH = CH(\varepsilon, X, Z^P, D)$$

We have denoted cross-hauling with *CH* and heterogeneity with ε. Both *CH* and ε are column vectors of dimension *n*. Our approach is based on estimating ε from equation (21). Before we can do so, we need to do some accounting. Let the total trade volume *V* be defined as the sum of (gross) exports and (gross) imports:

(22) $V = E + M$

The trade balance, or net export, is defined as the difference between exports and imports, which in turn is equal to domestic production minus domestic use:

(23) $B = E - M = X - Z^P - D$

This is, of course, the same equation which forms the cornerstone of the CB approach. Note that we are presuming an input-output table with indirect allocation of imports, because otherwise we do not have the information to calculate commodity balances. Using these two equations, we can write *M* and *E* as functions of *V* and *B*:

(24) $M = (V - B)/2$

(25) $E = (V - B)/2$

Furthermore, note that the trade volume can be written as the sum of the absolute value of the trade and the amount of cross-hauling:

(26) $V = |B| + CH$

If no cross-hauling is going on, equation (26) can only be fulfilled if either imports or exports (or both) are zero. In that case, our approach becomes equivalent to the classical CB approach outlined above. If, however, exports or imports are larger than zero, we know some amount of cross-hauling must be going on.

In order to be able to estimate ε, we must assume a specific functional form for (21). In order to keep things simple, we assume:

(27) $CH = \varepsilon(X + Z^S + D)$

Substituting this into (26) yields:

(28) $$V = |B| + \varepsilon(X + Z^S + D)$$

This equation can be solved for ε:

(29) $$\varepsilon = \frac{V - |B|}{X + Z^S + D}$$

The national input-output table contains data for all the variables on the right-hand side of (29). We can use these data to acquire an estimate of ε.

IV.3 Estimating the Regional Trade Pattern

In order to estimate the regional trade pattern, we first estimate regional production and consumption. Assuming that the only regional data we have is sectoral employment data, we can estimate the regional production of each sector with the following formula:

(30) $$X_i^R = \frac{L_i^R}{L_i^N} X^N$$

Equation (30) is based on the assumption that the regional labour productivity is equal to the national average for each sector.

Next, we estimate regional intermediate use. We calculate the matrix of national technical coefficients, A^X, as in equation (15) and assume that the region uses the same technology as the nation. Under this assumption, intermediate use can be estimated by:

(31) $$Z^{SR} = A^X X^R$$

Regional final use (excluding exports) is estimated by:

(32) $$D^R = \frac{L^R}{L^N} D^N$$

From equations (30), (31) and (32), we have estimates of X^R, Z^{SR} and D^R. Using those estimates in equation (23), we can acquire an estimate of the region's trade balance B^R. Then, equation (28) can be used to calculate an estimate of the region's trade volume V^R. Finally, equations (24) and (25) can be used to calculate estimates of regional imports M^R and exports E^R. Thus, the estimate of the regional trade pattern is complete.

The estimates calculated so far can then be used to complete the RIOT. The estimate of regional production X^R, along with the technical coefficients matrix A^X, can be used to calculate the entire matrix of interindustry transactions Z^R. By adding the estimates of regional production X^R and regional imports M^R, we get regional total supply S^R. Similarly, adding the estimates of regional intermediate use Z^S, regional final use D^R and regional exports E^R, we get regional total use U^R. The market clearing condition $S^R = U^R$ may then serve as a consistency check. Thus, the CHARM approach delivers a complete estimated RIOT.

Note that because *CH* is allowed to be larger than zero in the CHARM approach, the trade volume *V* tends to be larger than in the CB approach, which assumes that *CH* equals zero. Because of this, the output multipliers that are calculated from the Leontief inverse tend to be smaller. Thus, the overestimation of multipliers, which is an often-criticised feature of traditional nonsurvey methods, is meliorated with the CHARM approach.

IV.4 Application as Part of a Hybrid Approach

The CHARM method presented above is a pure nonsurvey method which requires exactly the same information as its cousins, the LQ and CB approaches, for implementation. Thus, it provides an alternative to traditional nonsurvey approaches. This means that it can also be used in the first phase of the top-down hybrid approach.

One possible drawback of the very simple CHARM method is that it relies on a good estimate of regional final use. In effect, we have assumed above that the regional final use vector is strictly proportional to its national counterpart. This is the same assumption which underlies implicitly the LQ and CB approaches.

However, if CHARM is implemented as part of a hybrid approach, its accuracy might be significantly increased by a better estimate of regional final use. Since Lahr (1993) suggests that the household sector should be one of the first to be surveyed anyway, no additional costs are incurred. Furthermore, disaggregated data on regional consumption by households, which accounts for the lion's share of final use, is generally available in many countries. Thus, a survey may not even be necessary, as reasonable estimates may be derived from existing databases.

CHARM can be applied only to tables with indirect allocation of imports, and consequently yields regional tables with indirect allocation. As pointed out above, LQ methods can only be applied to direct allocation tables. For some research questions, however, regional tables with indirect allocation are required. If these are to be constructed using LQ method, imports must be reallocated; a procedure which certainly does not improve the accuracy of the table. In those cases, CHARM may provide a very interesting alternative.

V Conclusion

It was realized by Richardson (1985), if not earlier, that the common regionalization methods are unsatisfying from a theoretical viewpoint, because they are based on four highly restrictive assumptions:

1) Equal labour productivity in the region and the nation
2) Equal consumption per employee in the region and the nation
3) If the region exports a commodity, it does not import the same commodity
4) The nation neither imports nor exports any given commodity in net terms

In this paper, we have presented the CHARM method, which provides an alternative to traditional nonsurvey methods and may be superior because it is based on less restrictive assumptions. Notably, assumptions 3 (no cross-hauling) and 4 (no net exports or imports at the national level) may be completely dropped. Assumption 1 (equal labour productivity) is necessary if the only regional data is on sectoral employment; if regional data on labour income by sector are available, assumption 1 may be dropped as well. Assumption 2 may be dropped as soon as data on regional consumption by households are available, which is often the case. Thus, in many instances CHARM may be applied without making the four restrictive assumptions criticised by Richardson [1985].

We conclude from this that CHARM provides an improvement over traditional nonsurvey methods, at least from a theoretical viewpoint. It remains to be seen whether its empirical performance can also compare. However, CHARM may not be directly comparable to LQ methods because the former can be applied only to IO tables with indirect allocation of imports, whereas the latter can be applied only to table with direct allocation. Therefore, it might make sense to use CHARM for the construction of regional tables with indirect allocation and still use traditional LQ methods for tables with direct allocation.

Like other nonsurvey methods, CHARM may also be applied as part of a top-down hybrid approach. Following a suggestion by West [1990], research on the construction of regional input-output tables seems to have focussed on developing other aspects of those hybrid approaches. Relatively little attention has been paid to advancing the state of the art of nonsurvey methods. However, West's suggestion seems to have sprung from the experience of working in Australia, where traditional nonsurvey methods perform reasonably well because cross-hauling is not a major issue. In more densely populated countries, where cross-hauling is significant, the development of nonsurvey methods which acknowledge the important role of cross-hauling is likely to significantly improve the performance of the hybrid approach. With this paper we hope to have made a small contribution to this end.

References

AUGURZKY, B. & NEUMANN, U. (2005) *Economic resources of senior citizens: Regional economic and fiscal effects of promoting the senior citizens economy in North Rhine-Westphalia (NRW).* MGSFF Ministerium für Gesundheit, Soziales, Frauen und Familie des Landes Nordrhein-Westfalen.

BONFIGLIO, A. (2005) *Can Non-survey Methods Substitute for Survey-based Models? A Performance Analysis of Indirect Techniques of Estimating I-O Coefficients and Multipliers.* Università Politecnica delle Marche, Dipartimento di economia. Quaderno di ricerca n. 230.

BOOMSMA, P. & OOSTERHAVEN, J. (1992) A double entry method for the construction of biregional input-output tables. *Journal of Regional Science,* 32, 269-284.

CONSAD RESEARCH CORPORATION (1967) *Regional Federal Procurement Study*. Office of Economic Research, U.S. Department of Commerce Contract 7-35211.

DIETZENBACHER, E. & VELÁZQUEZ, E. (2007) Analysing Andalusian Virtual Water Trade in an Input-Output Framework. *Regional Studies,* 41:2, 185-196.

EUROSTAT (1996) *European system of accounts: ESA 1995.* http://forum.europa.eu.int/irc/dsis/nfaccount/info/data/esa95/en/titelen.htm. April 17, 2007.

FLEGG, A. T., et al. (1995) On the Appropriate Use of Location Quotients in Generating Regional Input-Output Tables. *Regional Studies,* 29:6, 547-561.

GABRIEL, C. (2001) *Constructing Regionalized Input-Output Tables: A New Simple-to-use Method*. In PFÄHLER, W. (Ed.) *Regional Input-Output Analysis.* Baden-Baden, Nomos.

HARRIS, R. I. D. & LIU, A. (1998) Input-Output Modelling of the Urban and Regional Economy: The Importance of External Trade. *Regional Studies,* 32:9, 851-862.

ISARD, W. (1951) Interregional and Regional Input-Output Analysis: A Model of a Space-Economy. *Review of Economics and Statistics,* 33, 319-328.

ISARD, W. (1953) Regional Commodity Balances and Interregional Commodity Flows. *American Economic Review,* 43, 167-180.

ISSERMAN, A. M. (1980) Estimating Export Activity in a Regional Economy: A Theoretical and Empirical Analysis of Alternative Methods. *International Regional Science Review,* 5:2, 155-184.

JENSEN, R. C. (1978) Some accounting procedures and their effects on input-output multipliers. *Annals of Regional Science,* 12, 21-38.

JENSEN, R. C. (1980) The Concept of Accuracy in Regional Input-Output Models. *International Regional Science Review,* 5:2, 139-154.

JENSEN, R. C. (1990) Construction and Use of Regional Input-Output Models: Progress and Prospects. *International Regional Science Review,* 13:1&2, 9-25.

KLAASSEN, L. H. & VERSTER, A. C. P. (1974) *Kosten-batenanalyse in regionaal perspectief.* Groningen, Tjeenk Willink.

KRONENBERG, T. H. (2007) Derivative Construction of Regional Input-Output Tables under Limited Data Availability. STE Preprint 10/2007, Forschungszentrum Jülich.

LAHR, M. L. (1993) A review of the literature supporting the hybrid approach to constructing regional input-output models. *Economic Systems Research,* 5:3, 277-294.

LAHR, M. L. (2001a) Reconciling Domestication Techniques, the Notion of Reexports and Some Comments on Regional Accounting. *Economic Systems Research,* 13:2, 165-179.

LAHR, M. L. (2001b) *A Strategy for Producing Hybrid Regional Input-Output Tables*. In LAHR, M. L., et al. (Eds.) *Input-Output Analysis: Frontiers and Extensions.* London, Palgrave.

LEONTIEF, W. (1986) *Input-Output Economics.* New York and Oxford, Oxford University Press.

LLOP, M. (2007) Economic structure and pollution intensity within the environmental input-output framework. *Energy Policy,* 35:3410-3417.

MAKS, J. A. H. & OUDE WANSINK, M. J. (1998) Trade in a Regional Input-Output System. University of Maastricht, Eurecom Research Memorandum RM9901.

MIDMORE, P., et al. (2006) Assessing industry linkages using regional input-output tables. *Regional Studies,* 40:3, 329-343.

MILLER, R. E. & BLAIR, P. D. (1985) *Input-Output Analysis: foundations and extensions.* New Jersey, Prentice-Hall.

MOORE, F. T. & PETERSEN, J. W. (1955) Regional Analysis: An Interindustry Model of Utah. *Review of Economics and Statistics,* 37, 368-383.

MORRISON, W. I. & SMITH, P. (1974) Nonsurvey Input-Output Techniques at the Small Area Level: An Evaluation. *Journal of Regional Science,* 14:1, 1-14.

NORCLIFFE, G. D. (1983) Using location quotients to estimate the economic base and trade flows. *Regional Studies,* 17, 61-168.

PARK, S. H., et al. (1981) Error in regional nonsurvey input-output models: analytical and simulation results. *Journal of Regional Science,* 21, 321-37.

RICHARDSON, H. W. (1985) Input-output and economic base multipliers: Looking backward and forward. *Journal of Regional Science,* 25, 607-661.

RIDDINGTON, G., et al. (2006) Comparison of Gravity Model, Survey and Location Quotient-based Local Area Tables and Multipliers. *Regional Studies,* 40:9, 1069-1081.

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

ROUND, J. L. (1978) An interregional input-output approach to the evaluation of nonsurvey methods. *Journal of Regional Science,* 18, 179-194.

ROUND, J. L. (1983) Nonsurvey Techniques: A Critical Review of the Theory and Evidence. *International Regional Science Review,* 8:3, 189-212.

SCHAFFER, W. A. & CHU, K. (1969) Nonsurvey Techniques for Constructing Regional Interindustry Models. *Papers of the Regional Science Association,* 23, 83-101.

SCHINTKE, J. & STÄGLIN, R. (1988) *Important input costs in market transaction tables and production flow tables*. In CIASCHINI, M. (Ed.) *Input-Output Analysis: Currect Developments.* London, Chapman & Hall.

STEVENS, B. H., et al. (1983) A New Technique for the Construction of Non-Survey Regional Input-Output Models. *International Regional Science Review,* 8:3, 271-286.

STONE, R. (1961) *Input-Output and National Accounts.* Paris, Organization for European Economic Cooperation.

TOHMO, T. (2004) New Developments in the Use of Location Quotients to Estimate Regional Input-Output Coefficients and Multipliers. *Regional Studies,* 38:1, 43-54.

WEST, G. R. (1980) Generation of regional input-output tables (GRIT): An introspection. *Economic Analysis and Policy,* 10, 71-86.

WEST, G. R. (1981) An Efficient Approach to the Estimation of Regional Input-Output Multipliers. *Environment and Planning A,* 13, 857-867.

WEST, G. R. (1982) Sensitivity and key sector analysis in input-output models. *Australian Economic Papers,* 31, 857-867.

WEST, G. R. (1990) Regional Trade Estimation: A Hybrid Approach. *International Regional Science Review,* 13:1&2, 103-118.

Acknowledgement

This paper is part of the Integrated Project *Infrastructures and demographic development – Framework conditions and challenges in the context of sustainable development*. We gratefully acknowledge funding from the 'Impuls- und Vernetzungsfonds' of the Helmholtz-Gemeinschaft.

STE-Preprints 2007

01/2007 Martinsen, D., Krey, V., Markewitz, P.: Implications of High Energy Prices for Energy System and Emissions. - The Response from an Energy Model for Germany.

02/2007 Hake, J.-Fr., Fischer, W., Eich, R., Schlör, H.: Kernenergie im Kontext nachhaltiger Entwicklung. Aspekte der Proliferation, Entsorgung und Akzeptanz.

03/2007 Fischer, W.: www.InfrastrukturInternet-Cyberterror.Netzwerk. Analyse und Simulation strategischer Angriffe auf die kritische Infrastruktur Internet.

04/2007 Schlör, H., Hake, J.-F., Fischer, W.: How sustainable is the German Energy System? – A new concept for measuring sustainability.

05/2007 Markewitz, P., Vögele, S.: Entwicklungen und Trends in der Kraftwerkstechnik: Was können wir aus der Vergangenheit lernen?

06/2007 Krey, V., Markewitz, P., Vögele, S.: Energietransport und –verteilung.

07/2007 Nussbaum, S., Niemeyer, I., Canty, M.: Targeted Information Collection for Nuclear Verification: A Combination of Object-based Images Analysis and Pixel-Based Change Detection with Very High Resolution Satellite Data exemplified for Iranian Nuclear Sites.

08/2007 Autrusson, B., Richter, B.: Electronic Safeguards Seals.

09/2007 Rezniczek, A., Remagen, H., Richter, B., Stein, G., Blohm-Hieber, U., Nackaerts, H.: Experiences in Preparing the Implementation of the Additional Protocol in Germany.

10/2007 Kronenberg, T.: Derivative Construction of Regional Input-Output-Tables under Limited Data Availability.

11/2007 Nußbaum, S., Niemeyer, I.: Automated extraction of change information from multispectral satellite imagery.

12/2007 Nazarko, J., Schreiber, A., Kuckshinrichs, W., Zapp, P.: Environmental Analysis of the Coal-based Power Production with Amine-based Carbon Capture.

13/2007 Möller-Ühlken, K.: Demografische Determinanten der Energienachfrage der Haushalte in Deutschland.

STE-Research Reports 2007

01/2007 Markewitz, P., Vögele, S.: Schätzung der Wassernachfrage großer Kraftwerke im Elbeeinzugsgebiet

Systemforschung und Technologische Entwicklung im Forschungszentrum Jülich

Viele der im Brennpunkt des gesellschaftlichen Interesses stehenden Fragen lassen sich nur durch eine fachübergreifende Systemanalyse beantworten. Dabei sind häufig naturwissenschaftlich-technische, ökonomische und ökologische Subsysteme, die miteinander in Wechselwirkung stehen, gleichzeitig zu untersuchen. Die Programmgruppe Systemforschung und Technologische Entwicklung (STE) greift diesen Ansatz auf und konzentriert sich mit ihren Arbeiten auf Fragen zur langfristigen Ausrichtung der Energiewirtschaft, auf ausgewählte ökonomisch bzw. ökologisch relevante Stoffströme in Techno- und Geosphäre sowie auf elektronische Informationsverarbeitung und Kommunikation und dadurch verursachte Veränderungen in der Gesellschaft. Auf diesen Gebieten analysiert die STE die Folgen technischer Entwicklungen und erstellt wissenschaftliche Entscheidungshilfen für Politik und Wirtschaft. Grundlagen dafür sind die methodische Weiterentwicklung von Werkzeugen der Systemanalyse und ihre Anwendung sowie die Zusammenarbeit von Wissenschaftlern unterschiedlicher Fachrichtungen.

Systems Analysis and Technology Evaluation at the Research Centre Jülich

Many of the issues at the centre of public attention can only be dealt with by an interdisciplinary systems analysis. Scientific, economic and ecological subsystems which interact with each other often have to be investigated simultaneously. The Program Group Systems Analysis and Technology Evaluation (STE) takes up this approach and concentrates its work on issues concerning the long-term orientation of the energy economy, on selected economically or ecologically relevant material flows in the technosphere and geosphere as well as on electronic information processing and communications and the changes in society brought about by these technologies. In these fields, STE analyses the consequences of technical developments and provides scientific aids to decision making for politics and industry. This work is based on the further methodological development of systems analysis tools and their application as well as cooperation between scientists from different disciplines.

Leitung/Head: Jürgen-Friedrich Hake

Forschungszentrum Jülich
Institut für Energieforschung (IEF)
Systemforschung und Technologische Entwicklung (STE)
Leo Brandt Straße
52428 Jülich
Tel.: ..49-2461-61-6363
Fax: ..49-2461-61-2540,

Email: preprint-ste@fz-juelich.de
Web: www.fz-juelich.de/ste

STE Preprint

14
2007

Institut für Energieforschung (IEF)
Systemforschung und Technologische Entwicklung (STE)