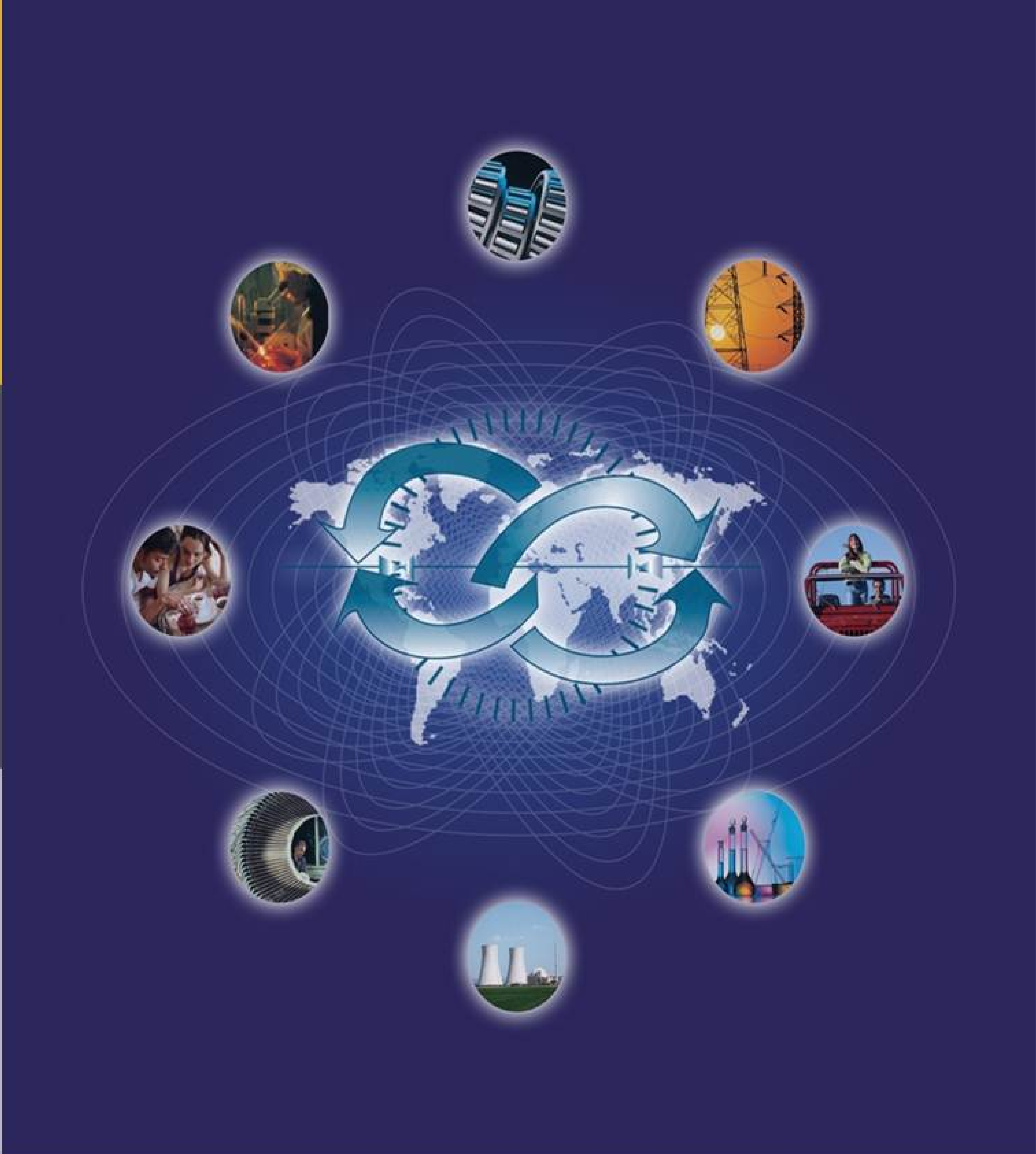


Mitglied der Helmholtz-Gemeinschaft

STE Preprint **01/2014**

Többen, Johannes

Construction of multi-regional supply-use tables: Experiences from Germany‘s federal states.

Institut für Energie- und Klimaforschung
Systemforschung und Technologische Entwicklung (IEK-STE)

JÜLICH
FORSCHUNGSZENTRUM

Construction of Multi-regional Supply-Use Tables: Experiences from Germany's Federal States

Johannes Többen[1)]

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

Executive Summary

Energy and climate policy in terms of national laws and incentive systems as well as international agreements are expected to induce an extensive structural change in many economies. The shape and implications of structural change caused by energy and climate policy has been subject of numerous studies, which concentrate either on Germany as a whole or on single regions. However, a multi-regional input-output analysis taking interdependencies across regions and sectors explicitly into account has not been undertaken so far.

This report is to illustrate the construction of the multi-regional supply-use table for Germany's federal states. The construction is based on a hybrid approach, which combines the estimation of a preliminary initial estimate using non-survey techniques with excessive amount of superior data from surveys on household consumption and industrial cost-structure. The estimation of interregional trade is based on transportations statistics. An important feature is the distinction of six types of income in the households sector. The consolidation initial estimate and superior data was conducted using the AISHA software package, a tool to construct (series) of large scale input-output tables.

Keywords

Multi-regional input-output, non-survey techniques, interregional trade, matrix balancing

Table of Contents

I Introduction

In recent years multi-regional input-output (MRIO) tables and models on the international level have received growing interested, due to an improved availability of economic accounts and trade data. The publication of international input-output databases such as EORA, EXIOPOL or WIOD has improved the opportunities to study international trade and supply chains in various fields of application at a high level of regional and sectoral disaggregation. In the light of global phenomena such as climate change and globalization multi-regional input-output models have become an important methodological framework for consumption-based accounting, for example of greenhouse gas emissions (see [Wiedmann, 2009] for a comprehensive review), water use [Daniels et al., 2011] or loss of biodiversity [Lenzen et al., 2012].

Originally, the theoretical basis of the interregional input-output model was developed by [Isard, 1951] for the subnational level. The so-called Isard-model is based on full information about regional as well as sectoral origin and destination of transactions. The term multi-regional input-output model originally refers to a simplified framework of the Isard-model developed by Chenery [1953 and Moses [1955, which only requires information about regional origin and destination of transactions along with information on regional technology and final consumption structures. Examples of subnational MRIO tables and applications in various fields such as (transport) infrastructure planning, regional development, environmental accounting or energy use include inter alia the Netherlands [Eding et al., 1999, Oosterhaven, 1981, Oosterhaven & Knaap, 2003], Indonesia [Hulu & Hewings, 1993], China [Liang et al., 2007], Japan [Yi et al., 2007], Australia [Gallego & Lenzen, 2009] and the US [Gordon et al., 2007, Polenske, 1980, Polenske & Levy, 1975]. Compared to recent developments of MRIOs on the international level, however, number of up-to-date subnational MRIO tables and applications falls short. Whereas for international MRIOs national input-output tables as well as information on trade flows between country pairs are available for most countries, such information is most often not available for regions that are part of a nation.

In the case of Germany a multi-regional input-output table has never been compiled yet, although due to country-characteristics and developments in the last two decades a MRIO would have been and would be a useful tool for supporting political decision-making: First of all, Germany's political system is based on strong federal elements, which means that important decisions in many fields on policy e.g. economic, environmental- or energy are to a large extend made by local governments. In order to prevent too strong regional divergence fiscal equalization scheme for the federal states was established in 1950 and has constitutional rank. In addition to this regional re-allocation scheme two other developments caused (at least implicit) regional reallocation policies. The first is the so-called solidarity surcharge, a mark-up on income tax, which was explicitly introduced to finance investments into infrastructure in the former socialist eastern part after the German reunification and therefore promote regional convergence. The second is Renewable Energies Act levy, a mark-up on electricity prices, from which feed-in tariffs for renewable energy power plants are financed.

Although seemingly aspatial policy, it can be expected that the German energy policy has strong effects on regional economic development, as locations of production and locations of operation of renewables are not evenly distributed across the country, but are clustered in some areas. Photovoltaic panels for example are predominantly produced in eastern regions and installed in southern regions, whereas the financing is charged from all consumers of electricity except of heavy industry. The analysis of regional economic effects of the German energy policy was the reason for the compilation of the multi-regional input-output table for Germany.

The lack of data caused the development of various non-survey methods such as the simple location-quotient and its many refinements such as CILQ or FLQ [Bonfiglio & Chelli, 2008, Flegg & Tohmo, 2011, Schaffer & Chu, 1969, Tohmo, 2004] or the commodity-balance method with its recent extension CHARM [Kronenberg, 2009]. These methods are based on only few basic data, e.g. as regional employment by industry and assumptions, such as regional invariance of technology, for generating single-regional input-output tables from national ones. The justification of the use of such methods is the concept of holistic accuracy [Jensen, 1980], which states that a reasonable accuracy of the results can be achieved, if at least most important intra- and interregional transactions can be estimated with high certainty. For the improvement of the nonsurvey table additional superior data can be introduced resulting in a hybrid approach (see Lahr [1993 for a review) that can be seen as a pragmatic compromise between costly and time-consuming survey-based and pure non-survey tables.

For the construction of a full MRIO the set of single-regional input-output tables have to be connected to each other via trade flows between regions. The approaches to accomplish this task range from using simple proportions (e.g. Bonet [2005) over models of spatial interaction, such as gravity or entropy-maximising approaches [Isard et al., 1998, Jensen-Butler & Madsen, 1996, Leontief, 1963, Wilson, 1970], to the use of regional transportation data [Gallego & Lenzen, 2009, Louhela, 2006] or survey information on region-specific sales or purchases of companies, i.e. of the wholesale sector [Eding et al., 1999, Piispala, 1999] .

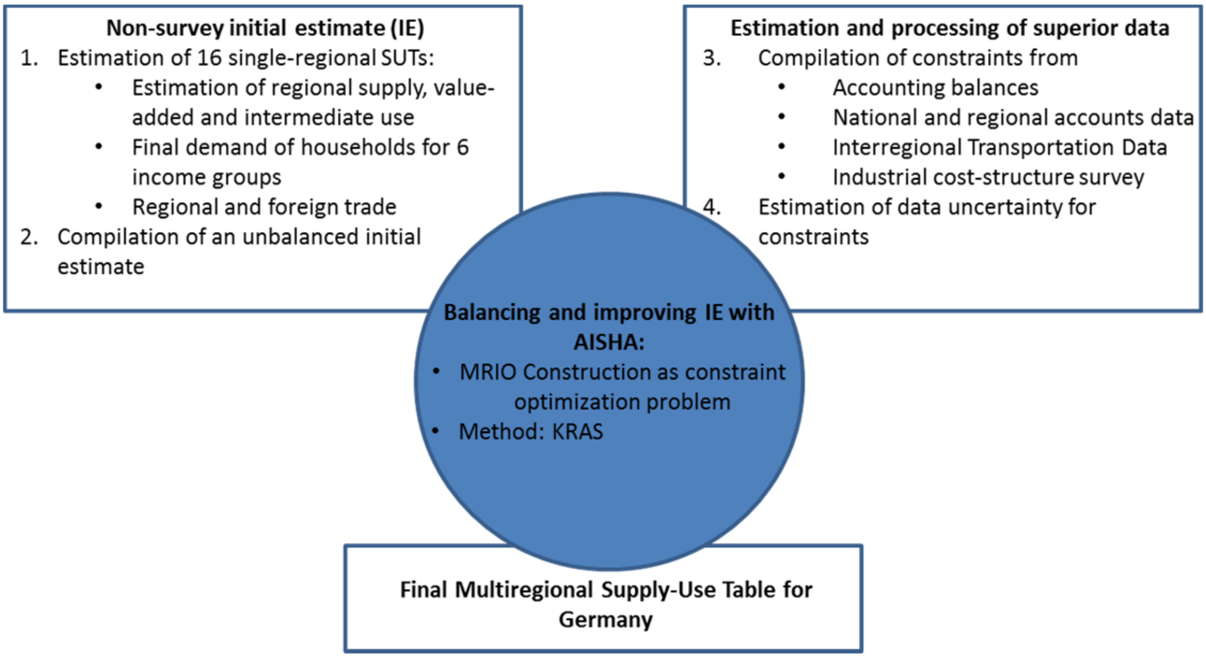
The resulting single- or multiregional tables are not usually balanced in the sense that economic accounting balances, e.g. output equals total use of products, are not fulfilled. A widely used method for balancing or updating tables is the RAS technique, which is based on the work of [Stone, 1962]. The result of RAS consists of a final table with minimal distance from the original one that satisfies constraints for row and column sums. It is important to note that the quality of final table crucially depends on the quality of the initial estimate and the methods used to generate it [Round, 1983]. The original RAS method was refined in several directions to deal with constraints on arbitrary sets of elements, e.g. for the introduction of additional data as constraints [Cole, 1992, Oosterhaven et al., 1986] , with negatives [Junius & Oosterhaven, 2003] or uncertainty of or conflicts in data [Dalgaard & Gysting, 2004, Lenzen et al., 2009] that are used for constraints.

Drawing on these experiences and advances the construction of the German multiregional supply-use table is carried out as a hybrid approach, whereby superior data is introduced during balancing process by making use of the tool AISHA [Geschke et al., 2011], whose core is the most recent RAS variant KRAS [Lenzen et al., 2009]. AISHA was developed at the University of Sydney for the compilation of EORA and offers a high degree of standardization and automation to the user. KRAS is able to deal with conflicting data on the basis of standard deviations of raw data, at the same time it allows the estimation and publication of reliability information of the final tablebased on standard error propagation [Lenzen et al., 2013]. The steps of the construction process are shown in figure 1 and are oriented to the requirements of AISHA:

1. 16 single-regional tables for the German federal states are constructed by regionalizing the national supply-use table using information from regional accounts. For the purpose of latter applications final demand of households is estimated from a household income and expenditure survey for six types of income-groups. Finally, interregional and foreign trade is estimated with a modified version of CHARM [Kronenberg, 2009].
2. An unbalanced initial estimate of the multi-regional table is constructed by connecting the single-regional tables via trade flows between region pairs. These trade flows are estimated on the basis regional shares in total interregional trade of products.
3. For the sake of consistency with national and regional accounts additional the national supply use-table as well as data on regional foreign trade, value-added, wages and output are used in addition to balancing constraints. In order to improve the quality of the table and to introduce the spatial dimension data on interregional transportation flows and information from an industrial cost-structure survey are used. The use of the data sources required several steps of estimation and processing.
4. For solving potential conflicts in external data, data uncertainty in terms of standard deviations is estimated for the constraints.

In following sections the supply-use framework, on which the German MRIO is based, is introduced for a single region and further extended to the case of many regions. Particular attention will be paid to input-output modeling based on supply-use tables and their relationship to symmetric input-output tables. Thereafter, the construction of the initial estimate is described in detail, whereby the need for modifying CHARM will be discussed and a refined version of the method will be developed. The following section deals with the estimation and processing of superior data for the improvement of the initial estimate as well as with the estimation of data uncertainty. Finally, the resulting final table will be discussed, which results from putting all these data are together by making use of AISHA.

Figure 1 Structure of the construction procedure



Source: own calculations IEK-STE 2013

II The Supply-Use Accounting Framework

The dimension of a symmetric input-output table can either be product-by-product or industry-by-industry. Product-by-product tables contain homogeneous products in the rows that are produced by homogenous branches (columns), so that it describes technology of production. An industry-by-industry table however consists of industry-output (row) and industries (column), whereby each transaction is considered as a bundle of different products. Due to this characteristic, industry-by-industry tables describe inter-industry relations rather than production technology. With respect to economic analysis product-by-product tables are considered to be the more appropriate type for applications related to effects of changes in technology, such as improvements in energy efficiency of steel production. However, industry-by-industry tables are considered to be better suited for the assessment of impacts on industries, such as employment effects of infrastructure projects [Eurostat/European Commission, 2008].

Unlike symmetric Input-Output tables, which are either of the industry-by-industry or product-by-product type, the Supply-Use framework incorporates both dimensions. The introduction of the Supply-Use accounting framework was largely induced by the fact, that industries often produce more than one product like secondary- or by-products. Due to differentiating between the production of products by industries and the use of products by industries and categories of final demand, heterogeneity of production can be accounted explicitly. For the purpose of input-output analysis supply-use tables can be transformed into industry-by-industry and product-by-product tables imposing assumptions on either production technology or market-shares. This framework therefore offers a lot of flexibility for later applications of the German multi-regional supply-use-table [Eurostat/European Commission, 2008].

There are further advantages for using supply-use tables instead of symmetric IO tables as the basis for models. When it comes to the construction for a single or for many regions on a sub-national level or when several a national tables are linked to an international table, data about trade flows are crucially important, in order to capture spatial spillover effects. The advantages of supply-use tables especially come into play during this construction process, as data on trade flows, consumption and production data are classified by products, whereas data about regional value added or wages are classified by industry (Madsen and Jensen-Butler 1999; [Bouwmeester & Oosterhaven, 2008]. Data can be more easily used for non-survey methods within a supply-use framework, as it incorporates both dimensions: products and industries [Oosterhaven, 1984]. The supply-use framework is also used by statistical offices for the collection of economic data and it is, therefore, argued by Madsen and Jensen-Butler [1999] that more accurate results could be achieved, since regional data can be used directly in contrast to the regionalization of square tables, which often require a re-classification of data. Hence, Oosterhaven [1984, p. 574] concludes that "instead of trying to force data into a square format, one may better present them in the format in which they are assembled of firms and consumers".

The Supply-Use framework has also advantages on theoretical ground: It is furthermore argued that models based on the supply-use framework are much closer related to theories of production, consumption and trade of commodities[Madsen & Jensen-Butler, 1999] In addition supply-use tables are considered to make implications of assumptions of models based on them easier to understand, in comparison models based on symmetric tables, which they call a "reduced form" [p. 278].

In the following subsections the supply-use framework is introduced and models, based on the transformation of supply-use tables into symmetric input-output tables under assumptions on market-shares or technology are discussed. Afterwards the supply-use framework is extended to the multi-regional level.

II.1 Supply-Use and Symmetric Input-Output Tables

Table 1 presents the basic structure of the Supply-Use framework according to the European System of Accounts [Eurostat, 1996], whereby the Supply table appears transposed and is then called Make table:

Table 1 Supply-use framework for a single region

	Product	**Industry**	**Final demand**	**Exports**	**Total**
Product		$\boldsymbol{U}$	$\boldsymbol{d}$	$\boldsymbol{e}$	$\boldsymbol{s}$
Industry	$\boldsymbol{V}$				$\boldsymbol{g}$
Value Added		$\boldsymbol{w}$			
Output	$\boldsymbol{p}$	$\boldsymbol{g}$			
Imports	$\boldsymbol{m}$				
Total	$\boldsymbol{s}$				

Source: own calculations IEK-STE 2013

The Supply-Use framework encompasses two fundamental sub-matrices: the use-matrix $\boldsymbol{U} = [u_{ji}]$ and the make-matrix $= [v_{ij}]$. Element u_{ji} indicates the amount of product *j* that is used by industry *i* as an intermediate input regardless of its origin, therefore this type of use-matrix is also known as total flow use-matrix in contrast to domestic flow use-matrices[1]. Val-

[1] Domestic flow use-matrices are not published by the federal statistical office. Also, use tables are valuated at purchaser prices and make tables at basic prices with a transition to purchaser prices.

ue added is arranged as the row vector $\boldsymbol{w} = [w_i]$ below the use-matrix and contains the sum primary inputs used by industry *i* for the production of its output as well as taxes less subsidies on products and production. The generic element of the make-matrix v_{ij} represents the domestic production of product *j* by industry *i*. Imports of product *j* are represented as the row vector $\boldsymbol{m} = [m_j]$ under the make-matrix. On the right hand side of use-matrix two column vectors $\boldsymbol{d} = [d_j]$ and $\boldsymbol{e} = [e_j]$ stand for domestic final use and exports of product *j*. ***d*** is the sum of final consumption expenditure by households, government and non-profit organization serving households (NPISH), as well as gross capital formation.

Based on these components it can be observed that two fundamental identities must hold for products and industries. Firstly, the total output by industry must be equal the sum of intermediate consumption and value-added by industry, that is:

$$g_i = \sum_j v_{ij} = \sum_j u_{ji} + w_j \qquad (1)$$

, where $\boldsymbol{g} = [g_i]$ is a generic element of the total output by industry vector. The second identity is the product supply-demand balance:

$$s_j = \sum_i v_{ij} + m_j = \sum_i u_{ji} + d_j + e_j \qquad (2)$$

Equation (2) states that the total supply of product *j* (domestic production and imports) must equal the total use of product *j* as an intermediate input or for final consumption. Written in matrix terms (1) and (2) appear as:

$$\boldsymbol{g} = \boldsymbol{V}\boldsymbol{i} = \boldsymbol{i}'\boldsymbol{U} + \boldsymbol{w} \qquad (1a)$$

$$\boldsymbol{s} = \boldsymbol{i}'\boldsymbol{V} + \boldsymbol{m} = \boldsymbol{U}\boldsymbol{i} + \boldsymbol{f} \qquad (2a)$$

, where $\boldsymbol{i}$and $\boldsymbol{i}'$indcate summation over rows and columns respectively and $\boldsymbol{f} = \boldsymbol{d} + \boldsymbol{e}$ denotes total final demand.

Similar to symmetric input-output table derivation of model requires the connection of purchases of intermediate inputs by industry with output by industry. Technical coefficients are then defined as:

$$\boldsymbol{B} = \boldsymbol{U}\widehat{\boldsymbol{g}}^{-1} \qquad (3)$$

$\boldsymbol{B}$contains “true” technical coefficients, as its elements b_{ij} account for intermediate consumption of domestic as well as imported goods. The hat indicates vectors written as diagonal matrices. Solving (3) for ***U*** and substituting into (2a) yields:

$$\boldsymbol{s} = \boldsymbol{B}\boldsymbol{g} + \boldsymbol{f} \qquad (4)$$

It is obviously not possible to generate a total requirements matrix from (4) without transforming either industry output into commodity output or vice versa by using only infor-

---

For sake of simplicity the use matrix is assumed to be valuated at basic prices, as the problem of valuation concepts is tackled below.

mation contained in the Make-matrix [Miller & Blair, 2009]. The conversion into either commodity- or industry output implies different assumptions.

The first alternative is to define column coefficients from the make-matrix:

$$\boldsymbol{D} = \boldsymbol{V}\hat{\boldsymbol{s}}^{-\mathbf{1}} <=> Ds = g \qquad (5).$$

D is the matrix of commodity supply proportions, whose elements $d_{ij} = v_{ij}/s_j$ denote the share of total product supply *j* that is produced by industry *i* or imported, assuming each product to be supplied in fixed proportions by domestic industries and the rest of the world. Hence, ***D*** is also known as the market shares matrix. Substituting (5) into (4) and solving for ***s*** yields a solution that expresses the relationship between final demand for domestic and imported products and total product supply:

$$\boldsymbol{s} = (\boldsymbol{I} - \boldsymbol{BD})^{-\mathbf{1}}\boldsymbol{f} \qquad (6).$$

In other words, (6) is equivalent to an input-output model based on a product-by-product table including domestic and imported products in the transaction matrix[2]. In many applications the effect on imported products needs to be separated from the effect on domestic production. Similarly, it is often favorable to only analyze the effect of final demand for domestic products on domestic output. Both purposes require the definition of a relationship between total and domestic flows. For product supply we have:

$$\boldsymbol{s} = \boldsymbol{p} + \boldsymbol{m} = \boldsymbol{p} + \widehat{\boldsymbol{\alpha}}\boldsymbol{s} <=> s = (\boldsymbol{I} - \widehat{\boldsymbol{\alpha}})^{-\mathbf{1}}\boldsymbol{p} \qquad (7a),$$

where$\widehat{\boldsymbol{\alpha}}$is a diagonal vector of import ratios.On the demand side it is assumed that there are fixed proportions of imports and domestic output in product supply for any type of use. For intermediate and final demand the separation is given as:

$$\boldsymbol{f} = \boldsymbol{f^N} + \widehat{\boldsymbol{\alpha}}\boldsymbol{f} <=> f = (\boldsymbol{I} - \widehat{\boldsymbol{\alpha}})^{-\mathbf{1}}\boldsymbol{f^N} \qquad (7b)$$

and

$$\boldsymbol{U} = \boldsymbol{U^N} + \widehat{\boldsymbol{\alpha}}\boldsymbol{U} <=> U = (\boldsymbol{I} - \widehat{\boldsymbol{\alpha}})^{-\mathbf{1}}\boldsymbol{U^N} \qquad (7c),$$

where the superscript *N* indicates the domestic origin of products. This import proportionality assumption is considered as one possible approach for separating total flow tables into domestic and import tables in the Eurostat manual [2008] and it is suggested to perform this separation on the most detailed level of disaggregation. Although in reality import shares for a product may well differ over the types of uses, empirical assessment by[Oosterhaven et al., 2008] showed that the impact of this assumption is small[3].

---

[2] This type of symmetric tables is based on the European System of Accounts[Eurostat, 1996], which is closely related to what is denoted as type A of import allocation in the UN handbook [1993] on national accounts .

[3] Some authors such as [Jackson, 1998] and [Lahr, 2001] exclude exports from the import-proportionality assumption or apply different proportions on them, as the share of import in ex-

Using (6) and (7a) and solving for domestic output yields a product-by-product solution describing the impact of shifts in total final demand domestic product output:

$$\boldsymbol{p} = (\boldsymbol{I} - \widehat{\boldsymbol{\alpha}})(\boldsymbol{I} - \boldsymbol{BD})^{-1}\boldsymbol{f} \qquad (8)$$

If (7b) is additionally substituted into (8) a product-by-product solution is achieved that describes the effect final demand for domestic products on domestic product output:

$$\boldsymbol{p} = (\boldsymbol{I} - \widehat{\boldsymbol{\alpha}})(\boldsymbol{I} - \boldsymbol{BD})^{-1}(\boldsymbol{1} - \widehat{\boldsymbol{\alpha}})^{-1}\boldsymbol{f}^{\boldsymbol{N}} \qquad (9)$$

In the case that one wants to achieve a model that is equivalent to an input-output model based on symmetric industry-by-industry tables (4) is pre-multiplied by $\boldsymbol{D}$,(5) is substituted and solved for industry-output:

$$\boldsymbol{g} = (\boldsymbol{I} - \boldsymbol{DB})^{-1}\boldsymbol{Df} \qquad (10).$$

As the market-share matrix $\boldsymbol{D}$represents the contribution of each industry to total supply of products (domestic output and imports), pre-multiplication with $\boldsymbol{D}$transforms the effect of final demand on product supply to effects on domestic industries. Also, final demand for products is transformed into final demand for output of domestic industries[Sargento, 2009].

The second alternative is based on the matrix of row coefficients derived from the supply-matrix:

$$\boldsymbol{C} = \boldsymbol{V}'\widehat{\boldsymbol{g}}^{-1} <=> g = \boldsymbol{C}^{-1}\boldsymbol{p} \qquad (11).$$

The elements of $\boldsymbol{C}$, $c_{\boldsymbol{ij}} = v_{ij}/g_i$, denote the share of commodity *j* in the output of industry *i,* assuming that the proportions of different products in the output of each industry are fixed. ***C*** is, therefore, also called the product mix matrix. Similar (5) for the first alternative (11) can be used to derive models that are equivalent to input-output models based on product-by-product ore industry-by-industry tables. Substituting (11) into (4) yields:

$$\boldsymbol{s} = \boldsymbol{BC}^{-1}\boldsymbol{p} + \boldsymbol{f} \qquad (12).$$

Equation (12) can, then, be used along with import ratios defined in (7a) to derive models that describe the effect of total final demand on either product supply or domestic output:

$$\boldsymbol{p} = [(\boldsymbol{I} - \widehat{\boldsymbol{\alpha}})^{-1} - \boldsymbol{BC}^{-1}]^{-1}\boldsymbol{f} \qquad (13)$$

and

$$\boldsymbol{s} = [\boldsymbol{I} - \boldsymbol{BC}^{-1}(\boldsymbol{I} - \widehat{\boldsymbol{\alpha}})]^{-1}\boldsymbol{f} \qquad (14)$$

For the assessment of shift in final demand for domestic products (7b) may be substituted into (13) analogously to (9):

$$\boldsymbol{p} = [(\boldsymbol{I} - \widehat{\boldsymbol{\alpha}})^{-1} - \boldsymbol{BC}^{-1}]^{-1}(\boldsymbol{1} - \widehat{\boldsymbol{\alpha}})^{-1}\boldsymbol{f}^{\boldsymbol{N}} \qquad (15).$$

---

ported products (re-exports) often differs considerably from other types of use. The role of re-exports is discussed more deeply in section *III.1.3.*

Finally, it is also possibly to derive a model equivalent to those based on symmetric industry-by-industry tables. Substituting (7a) into (4), pre-multiplying with $\boldsymbol{C}^{-1}$ and solving for industry output yields:

$$\boldsymbol{g} = [(\boldsymbol{I} - \widehat{\boldsymbol{\alpha}})^{-1} - \boldsymbol{C}^{-1}\boldsymbol{B}]^{-1}\boldsymbol{C}^{-1}(\boldsymbol{1} - \widehat{\boldsymbol{\alpha}})^{-1}\boldsymbol{f}^{N} \quad (16).$$

On analogy to (10) the inverse product mix matrix is used to transfer final demand for domestic products into final demand directed to industries[Sargento, 2009].

As mentioned before, product-by-product tables require stating assumptions on production technology: Deriving product-by-product tables using (5) requires the assumption that all products produced by an industry to have the same input structure, so that each industry uses one characteristic technology. This assumption is therefore called Industry Technology Assumption (ITA). Making use of (11), however, implies that each product is produced with the same input structure irrespective of the industry that produces it, so that each product has one characteristic technology. For this reason the second assumption is known as Commodity Technology Assumption (CTA) [Miller & Blair, 2009, Sargento, 2009].

There has been a long lasting debate over pros and cons of both assumptions in the literature, but this controversial mostly concerns the case where the number of industries and products is equal (*n*=*m*). The reason simply is that there is no well-defined inverse of $\boldsymbol{C}$ (8a and 8b) in the case of rectangular tables ($n \neq m$), thus generation of total requirements matrices either requires aggregating commodities or industries until *n*=*m*, which goes along with a loss of information, or one has to choose an industry technology model[Miller & Blair, 2009]. Aside from models based on rectangular tables no consensus that prefers one of the assumptions could have been achieved yet. Jansen and ten Raa [1990] for example state four properties of the product-by-product direct requirement matrix $\boldsymbol{A}(\boldsymbol{U}, \boldsymbol{V})$that should be matched when deriving them from Supply-Use tables and prove that only the commodity-technology assumption satisfies all four criteria. The first property is the material balance meaning that total intermediate use of product *j* in the use table has to be equal to symmetric table. Financial balance, the second property, requires total revenue of product *j* to be equal to the sum of material cost and value added. Thirdly, price invariance implies that$\boldsymbol{A}(\boldsymbol{U}, \boldsymbol{V})$is unaffected fromthe choice of a base year if $\boldsymbol{U}$ and $\boldsymbol{V}$ are tables in constant prices. And finally, technical coefficients are expected to be invariant with regard to scaling factors. In contrast to the CTA ITA only fulfills the assumption of material balance.

On the other hand, the major drawback of the CTA is that it frequently produces negative transactions and thus negative elements in the direct requirement matrix when applied to real world input-output data, albeit most of them are small. The occurrence of negatives, which was first published by[van Rijckeghem, 1967], is obviously economically unreasonable. This problem is often seen as a result of errors in data or heterogeneous classification. However, De Mesnard [2011]argues that negatives arise systematically by proving that negatives in inverse supply-matrices derived from symmetric product-by-product tables occur even if the symmetric table contains no negatives. In a less recent paperhe argues that even if negatives do not occur, the commodity-technology assumption should be rejected, concluding that "either IBT [ITA] is adopted but violates Koop Jansen and Ten Raa's axioms [1990] or CBT [CTA] is chosen but one must convert it into a poor and implausible supply-driven model" [2004] p. 127. Almon [2000], however, finds the industry technology assumption "highly implausible" (citing the *System of National Accounts 1993*), because companies often produce secondary products that differ completely from the main product in terms of input structure (e.g. car manufacturers usually also supply financial services to customers). He also presents an algorithm that avoids negative values and was successfully used by the INFORUM group since 1967.

Industry-by-industry tables, on the other hand, do not require assumptions on technology to be made, although in literature the term "technology" is still frequently used. Eurostat finds the technology terminology misleading when applied to industry-by industry tables (as it was done in the SNA 1968) and differentiates between technology assumptions (industry and commodity based) that are used to derive product-by-product symmetric tables and sales structure assumptions that are used to derive industry-by-industry tables. Sales structure assumption can differentiate into fixed industry sales and fixed product sales meaning that either each industry output or each product has a specific sales structure. The assumption of fixed industry sales structure is also prone to negatives in the direct requirements matrix (the procedure also makes use of the inversion of $\boldsymbol{C}$) and [Mesnard, 2011] concludes that his arguments against CTA also apply to the assumption of fixed industry sales structure, despite of being rather implausible. In contrast to industry-sales structures Eurostat [2008]argues that the fixed product-sales structure assumption is widely used in statistical offices for the transmission to industry-by-industry tables, as it permits the occurrence of negative values. It is, furthermore, argued that overall sales structure in a row is actually not an assumption, since it is observed in the data.

Considering the application on regional structural change issues and advantages of simplicity and clarity compared to deriving product-by-product models based on commodity-technology assumptions, it was decided to focus on industry-by-industry models using fixed product sales-structure.

II.2 The Family of Multi-Regional Supply-Use Accounting Frameworks

In the construction of single and many-region input-output tables symmetric tables (industry-by-industry or product-by-product tables) of domestic flows have been predominant formats for a long time[4], but in recent years a growing interest in compiling rectangular Supply-Use tables for a single or many regions from national ones can be observed. Early attempts to extend the supply-use framework to multi- and interregional cases, analogue to symmetric models such as those developed by Chenery and Moses ([1953] and [1955]) or Isard[1951], were given by Hoffman and Kent [1974] and, in greater detail, by Oosterhaven [1984], who developed a whole family of different multi- and interregional supply-use ac-

[4]e.g. Miller and Blair's (2009) comprehensive textbook only deals with regional- and multiregional symmetric tables

counting frameworks. In a recent publication Jackson and Schwarm [2011] pick up this issue complaining a gap in literature on conceptual aspects of multi-regional supply-use tables.

The simplest possible variant of multi-regional supply-use frameworks, in which only information about the spatial origin and destination of trade flows is given, is presented in table 2 for the case of two regions $q = \{r, s\}$. In comparison to the single regional framework inTable 1$\boldsymbol{V}$is disaggregated into two regional supply tables $\boldsymbol{V^q} = [v_{ij}^q]$, whose generic element denote the amount of product *j* that is product by industry i in region q. Column sums $\boldsymbol{iV^q} = \boldsymbol{p^q}$ denote the total regional output by product. Imports of products are shown separated into their regional origin: $\boldsymbol{t^{rs}} = [t_j^{rs}]$ denotes the imports of product *j* of region s from region *r* and $\boldsymbol{m^s} = [m_j^1]$ denotes the imports of product *j* of region *s* from the rest of the world (row). The sum of domestic output and imports from other regions and the rest of world delivers the total supply of products within a region: $\boldsymbol{s^s} = [s_j^s]$.

Table 2 Multi-regional supply-use table

		Region r		Region s		Final Demand	Exports (region r)	Exports (region s)	Exports	Total
		Product	**Industry**	**Product**	**Industry**					
Region r	**Product**		$\boldsymbol{U^r}$			$\boldsymbol{d^r}$	**0**	$\boldsymbol{t^{rs}}$	$\boldsymbol{e^{r\,row}}$	$\boldsymbol{s^r}$
	Industry	$\boldsymbol{V^r}$								$\boldsymbol{g^r}$
Region s	**Product**				$\boldsymbol{U^s}$	$\boldsymbol{d^s}$	$\boldsymbol{t^{sr}}$	**0**	$\boldsymbol{e^{s\,row}}$	$\boldsymbol{s^s}$
	Industry			$\boldsymbol{V^s}$						$\boldsymbol{g^s}$
Value Added			$\boldsymbol{w^r}$		$\boldsymbol{w^s}$					
Output		$\boldsymbol{p^r}$	$\boldsymbol{g^r}$	$\boldsymbol{p^s}$	$\boldsymbol{g^s}$					
Imports (region r)		**0**		$\boldsymbol{t^{rs}}$						
Imports (region s)		$\boldsymbol{t^{sr}}$		**0**						
Imports (row)		$\boldsymbol{m^{row\,r}}$		$\boldsymbol{m^{row\,s}}$						
Total		$\boldsymbol{s^r}$		$\boldsymbol{s^s}$						

Source: own calculations IEK-STE 2013

The national use-table is also disaggregated into two regional tables: $\boldsymbol{U^q} = [u_{ji}^q]$ is the generic element regional use table and denotes the amount of product *j* that is used by industry i in region q as an intermediate input. Total final use of product *j* in region q is separated into domestic final use $\boldsymbol{d^q} = [d_j^q]$, exports to other regions $\boldsymbol{t^{sr}} = [t_j^{sr}]$ and to the rest of the world$\boldsymbol{e^q} = [e_j^q]$. Value added of industry *i* in region q is indicated by the row vector $\boldsymbol{w^q} = [w_i^q]$. Output of industry i in region q $\boldsymbol{g^q} = [g_i^q]$is then the sum of total intermediate consumption by industry $\boldsymbol{iU^q}$ and value added.

Within this framework following accounting identities must hold: The first one is the product supply-demand balance, which states that total supply of products has to equal to total use of products within a region:

$$\boldsymbol{s^q} = \boldsymbol{p^q} + \boldsymbol{t^{\cdot q}} + \boldsymbol{m^q} = \boldsymbol{iU^q} + \boldsymbol{d^q} + \boldsymbol{t^{q\cdot}} + \boldsymbol{e^q} \quad (17a).$$

The second identity states that total output by industry has to be equal to the sum of intermediate and primary inputs (value added) and is thus called industry input-output balance:

$$\boldsymbol{g^q} = \boldsymbol{iV^q} = \boldsymbol{U^q i} + \boldsymbol{w^q} \quad (18)$$

By making use of partitionated matrices (17) and (18) appear as:

$$\boldsymbol{s} = \boldsymbol{p} + \boldsymbol{iT} + \boldsymbol{m} = \boldsymbol{iU} + \boldsymbol{d} + \boldsymbol{iT'} + \boldsymbol{e} \quad (17b)$$

$$\boldsymbol{g} = \boldsymbol{iV} = \boldsymbol{Ui} + \boldsymbol{w} \quad (18b).$$

Whereby the matrices have the following form:

$$\boldsymbol{s} = \begin{pmatrix}\boldsymbol{s^r}\\ \boldsymbol{s^s}\end{pmatrix}; \boldsymbol{g} = \begin{pmatrix}\boldsymbol{g^r}\\ \boldsymbol{g^s}\end{pmatrix} \boldsymbol{p'} = \begin{pmatrix}\boldsymbol{p^r}\\ \boldsymbol{p^s}\end{pmatrix}; \boldsymbol{m'} = \begin{pmatrix}\boldsymbol{m^r}\\ \boldsymbol{m^s}\end{pmatrix}; \boldsymbol{d} = \begin{pmatrix}\boldsymbol{d^r}\\ \boldsymbol{d^s}\end{pmatrix}; \boldsymbol{e} = \begin{pmatrix}\boldsymbol{e^r}\\ \boldsymbol{e^s}\end{pmatrix}; \boldsymbol{w'} = \begin{pmatrix}\boldsymbol{w^r}\\ \boldsymbol{w^s}\end{pmatrix};$$

$$\widehat{\boldsymbol{T}} = \begin{pmatrix}\boldsymbol{0} & \boldsymbol{t^{rs}}\\ \boldsymbol{t^{sr}} & \boldsymbol{0}\end{pmatrix}; \widehat{\boldsymbol{U}} = \begin{pmatrix}\boldsymbol{U^r} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{U^s}\end{pmatrix}; \widehat{\boldsymbol{V}} = \begin{pmatrix}\boldsymbol{V^r} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{V^s}\end{pmatrix}$$

If one wants to derive an input-output model that describes the relationship between regional final demand and regional product- or industry output respectively, trade coefficients need to be defined that transform aspatial regional purchases to their origin of production:

$$\widehat{\boldsymbol{C}} = (\widehat{\boldsymbol{p}} + \widehat{\boldsymbol{T}})\widehat{\boldsymbol{s}}^{-1} \rightarrow \boldsymbol{g} = \boldsymbol{s}\widehat{\boldsymbol{C}} \quad (18).$$

The elements of $\widehat{\boldsymbol{C}} c_j^{qq}$ denote the share of a region in total supply of product j in each region. For products, that are also imported from the rest of the world, column sums of $\widehat{\boldsymbol{C}}$will be smaller than unity and will, therefore, separate total regional purchases (intraregional + imports from other regional + imports from the rest of world) into purchases from a specific region excluding imports from the rest of the world. Introducing $\widehat{\boldsymbol{C}}$along with technical coefficients $\boldsymbol{B}$ and sector-shares$\overline{\boldsymbol{D}}$and assuming equal average import propensity of all types of use (17b) may be rewritten as:

$$\boldsymbol{p} \equiv \widehat{\boldsymbol{C}}(\boldsymbol{B}\overline{\boldsymbol{D}}\boldsymbol{p} + \boldsymbol{f}) \quad (19)$$

The difference between the matrix $\overline{\boldsymbol{D}} = \boldsymbol{V}\widehat{\boldsymbol{p}}^{-1}$sector-shares matrix and the market-shares matrix $\boldsymbol{D}$is that sector shares do not imports into account. The reason for using sector-shares is that import from the rest of the world are already excluded by using $\widehat{\boldsymbol{C}}$, so that using market-shares would result in double-counting of foreign imports. Thereby, product $\widehat{\boldsymbol{C}}\boldsymbol{B}\overline{\boldsymbol{D}}\boldsymbol{p}$ as following properties: The sector-shares matrix $\overline{\boldsymbol{D}}$allocates regional product output $\boldsymbol{p}$ to regional industries according to their sector-shares and determines, therefore, regional output by industry. Regional industry-output further determines intermediate consumption of products by industries according to the technical coefficients matrix $\boldsymbol{B}$assuming industry technology. Thereafter, regional intermediate consumption is allocated to the regional origin of products and purchases from the rest of the world are excluded. Solving (19) for product output yields the solution of a multi-regional input-output model that describes the relationship between regional final demand and regional product output:

$$\boldsymbol{p} = (\boldsymbol{I} - \widehat{\boldsymbol{C}}\boldsymbol{B}\overline{\boldsymbol{D}})^{-1}\widehat{\boldsymbol{C}}\boldsymbol{f} \qquad (20).$$

The corresponding industry-by-industry model is derived as follows:

$$\boldsymbol{g} \equiv \overline{\boldsymbol{D}}\widehat{\boldsymbol{C}}(\boldsymbol{B}\boldsymbol{g} + \boldsymbol{f}) \rightarrow \boldsymbol{g} = (\boldsymbol{I} - \overline{\boldsymbol{D}}\widehat{\boldsymbol{C}}\boldsymbol{B})^{-1}\overline{\boldsymbol{D}}\widehat{\boldsymbol{C}}\boldsymbol{f} \qquad (21)$$

In this case pre-multiplying final demand by $\overline{\boldsymbol{D}}\widehat{\boldsymbol{C}}$, first, transfers final demand for products to purchases from a specific region and further to specific industries that deliver that products. This model is analogous to the Chenery-Moses-model for symmetric multi-regional tables in the sense that information about the spatial origin and destination of products is used to derive a matrix of (column) trade coefficients, which distributes aspatial purchases of industries and categories of (domestic) final demand within a region to different regional sources according to their market share in product supply [Oosterhaven, 1984].

If additional information about the destination of regional imports in terms of industry and categories of final demand are, given following interregional accounting framework can be constructed, whose corresponding model is analogue to Oosterhaven's [1984] purchase-only interregional and the use-regionalized model respectively [Jackson & Schwarm, 2011].Table 3 presents this framework.

Table 3 Use-regionalized inter-regional framework

		Region r			Region s			Exports (row)	Total
		Product	**Industry**	**Final Demand**	**Product**	**Industry**	**Final Demand**		
Region r	**Product**		$\boldsymbol{U^{rr}}$	$\boldsymbol{d^{rr}}$		$\boldsymbol{U^{rs}}$	$\boldsymbol{d^{rs}}$	$\boldsymbol{e^{r}}$	$\boldsymbol{p^{r}}$
	Industry	$\boldsymbol{V^{r}}$							$\boldsymbol{g^{r}}$
Region s	**Product**		$\boldsymbol{U^{sr}}$	$\boldsymbol{d^{sr}}$		$\boldsymbol{U^{ss}}$	$\boldsymbol{d^{ss}}$	$\boldsymbol{e^{s}}$	$\boldsymbol{p^{s}}$
	Industry				$\boldsymbol{V^{s}}$				$\boldsymbol{g^{s}}$
Imports (row)			$\boldsymbol{U^{row\ r}}$	$\boldsymbol{d^{row\ r}}$		$\boldsymbol{U^{row\ s}}$	$\boldsymbol{d^{row\ s}}$		
Value Added			$\boldsymbol{w^{r}}$			$\boldsymbol{w^{s}}$			
Output		$\boldsymbol{p^{r}}$	$\boldsymbol{g^{r}}$		$\boldsymbol{p^{s}}$	$\boldsymbol{g^{s}}$			

Source: own calculations IEK-STE 2013

The difference between Table 4 and the use-regionalized accounting framework is that the domestic demand is that intermediate and (domestic) final demands are now separated according to their spatial origin. $\boldsymbol{U^{qq}} = [u_{ji}^{qq}]$ denotes the demand of industry *i* in region q for product *j* produced in region *q* and $\boldsymbol{d^{qq}} = [d_{i}^{qq}]$ denotes the (domestic) final demand in region q for product *j* produced in region q. Hence, the link to the multi-regional table is the following: $\widehat{\boldsymbol{C}}\widehat{\boldsymbol{U}}$is an approximation of$\widetilde{\boldsymbol{U}} = [\boldsymbol{U^{qq}}]$ and $\widehat{\boldsymbol{C}}\boldsymbol{d}$ is an approximation of $\boldsymbol{F} = [\widehat{\boldsymbol{f}}^{\boldsymbol{qq}}]$ by assuming equal average import propensity of industries and categories of final demand (including exports to the rest of the world) $\boldsymbol{U^{rowq}}$ and $\boldsymbol{d^{row\,q}}$ are purchases from the rest of world for intermediate and final use in region q respectively. For the development of an interregional input-output model interregional input coefficients are derived from $\widetilde{\boldsymbol{U}}$as $\widetilde{\boldsymbol{B}} = \widetilde{\boldsymbol{U}}\widehat{\boldsymbol{g}}^{-\mathbf{1}}$. Each element of $\widetilde{\boldsymbol{B}}$denotes the share of product j produced in region q that is used in the own or the other region by industry i for intermediate consumption. The product-by-product model is then derived as follows:

$$\boldsymbol{p} \equiv \widetilde{\boldsymbol{B}}\overline{\boldsymbol{D}}\boldsymbol{p} + \boldsymbol{F}\boldsymbol{i} \rightarrow \boldsymbol{p} = (\boldsymbol{I} - \widetilde{\boldsymbol{B}}\overline{\boldsymbol{D}})^{-\mathbf{1}}\boldsymbol{F} \qquad (22)$$

In this case final demand for products from a specific region is first distributed to industries according to their sector-shares $\overline{\boldsymbol{D}}$and afterwards transferred into corresponding intermediate consumption of products from a specific region. Similar to $\widehat{\boldsymbol{C}}\boldsymbol{B}$in the multiregional model imports from the rest of the world are excluded. The corresponding industry-by-industry model can be derived as follows:

$$\boldsymbol{g} \equiv \overline{\boldsymbol{D}}(\widetilde{\boldsymbol{B}}\boldsymbol{g} + \boldsymbol{F}) \rightarrow \boldsymbol{g} = (\boldsymbol{I} - \overline{\boldsymbol{D}}\widetilde{\boldsymbol{B}})^{-\mathbf{1}}\overline{\boldsymbol{D}}\boldsymbol{F} \qquad (23)$$

In the interregional industry-by-industry model based on the use-regionalized framework final demand for products is at first allocated to industry output according to sector-shares. Corresponding intermediate purchases of products from a specific region is, as a second step, allocated to industries using sector-shares again [Jackson & Schwarm, 2011, Oosterhaven, 1984, Sargento, 2009] .

The third possible accounting framework is presented in Table 4, namely the make-regionalized [Jackson & Schwarm, 2011] or sales-only inter-regional table [Oosterhaven, 1984].

On contrary to use-regionalized framework, that requires additional information about the destination of regional imports in terms of industry and categories of final demand, information about the industrial origin is needed in addition to information of spatial origin and destination of trade flows. $\boldsymbol{V^{qq}} = [v_{ij}^{qq}]$ denotes the amount of product *j* produced by industry *i* in region q shipped to the same or to another region. Within this framework the technological coefficient matrix $\widehat{\boldsymbol{B}}$is used to distribute total intermediate demand for products in a region to the purchasing industries. The further distribution of these aspatial purchases to regions and industries of origin by means of a spatial market-share matrix $\widetilde{\boldsymbol{D}} = \widetilde{\boldsymbol{V}}\widehat{\boldsymbol{s}}^{-\mathbf{1}}$. The corresponding industry-by-industry interregional models can be derived as:

$$\boldsymbol{g} \equiv \widetilde{\boldsymbol{D}}(\boldsymbol{B}\boldsymbol{g} + \boldsymbol{f}) \rightarrow \boldsymbol{g} = \left(\boldsymbol{I} - \widetilde{\boldsymbol{D}}\boldsymbol{B}\right)^{-\mathbf{1}}\widetilde{\boldsymbol{D}}\boldsymbol{f} \qquad (24).$$

In this case total final demand is multiplied by the interregional market-shares matrix $\widetilde{\boldsymbol{D}}$and therefore transformed into final demand for output of industries in the respective regions, whereby final demand for imports from the rest of the world are excluded. Next, final demand for regional industry output is linked to intermediate consumption of domestic and imported products via the matrix of technical coefficients $\boldsymbol{B}$.

Table 4 make-regionalized inter-regional table

		Region r		Region s		Final Demand	Exports (region r)	Exports (region s)	Exports (row)	Total
		Product	**Industry**	**Product**	**Industry**					
Region r	**Product**		$\boldsymbol{U}^r$			$\boldsymbol{d}^r$	**0**	$\boldsymbol{t}^{rs}$	$\boldsymbol{e}^{r\,row}$	$\boldsymbol{s}^r$
	Industry	$\boldsymbol{V}^{rr}$		$\boldsymbol{V}^{rs}$						$\boldsymbol{g}^r$
Region s	**Product**				$\boldsymbol{U}^s$	$\boldsymbol{d}^s$	$\boldsymbol{t}^{sr}$	**0**	$\boldsymbol{e}^{s\,row}$	$\boldsymbol{s}^s$
	Industry	$\boldsymbol{V}^{sr}$		$\boldsymbol{V}^{ss}$						$\boldsymbol{g}^s$
Value Added			$\boldsymbol{w}^r$		$\boldsymbol{w}^s$					
Imports (row)		$\boldsymbol{m}^r$		$\boldsymbol{m}^s$						
Total		$\boldsymbol{s}^r$	$\boldsymbol{g}^r$	$\boldsymbol{s}^s$	$\boldsymbol{g}^s$					

Source: own calculations IEK-STE 2013

Thereafter, aspatial intermediate consumption is transformed into intermediate demand for industry-output of the regions (again excluding imports from the rest of the world) by multiplying with $\widetilde{\boldsymbol{D}}$ ([Jackson & Schwarm, 2011];[Oosterhaven, 1984]). The corresponding model for the product-by-product dimension in found by pre-multiplying (24) with sector-shares matrix $\overline{\boldsymbol{D}}$in order to relate effects of final demand on industry-output to regional output of products

$$\boldsymbol{p} = \overline{\boldsymbol{D}}\left(\boldsymbol{I} - \widetilde{\boldsymbol{D}}\boldsymbol{B}\right)^{-\mathbf{1}}\widetilde{\boldsymbol{D}}\boldsymbol{f} \qquad (25).$$

Finally, it is also possible to use both information on industrial as well as regional origin and destination of products in an interregional accounting framework. Table 5 presents such a framework. In contrast to the aforementioned accounting frameworks information about industrial origin and destination is necessary in addition to information on regional trade flows. For this reason regional use-tables as well as regional supply tables can be split up into regional origin of products used and regional destination of products supplied by an industry. From this frame work an input-output model with product-by-product dimension may be derived, by defining a matrix of region specific sector-shares $\widetilde{\widetilde{\boldsymbol{D}}} = \widetilde{\boldsymbol{V}}\widehat{\boldsymbol{p}}^{-\mathbf{1}}$:

$$\boldsymbol{p} \equiv \widetilde{\boldsymbol{U}}\widetilde{\widetilde{\boldsymbol{D}}}\boldsymbol{p} + \boldsymbol{F} \rightarrow \boldsymbol{p} = \left(\boldsymbol{I} - \widetilde{\boldsymbol{U}}\widetilde{\widetilde{\boldsymbol{D}}}\right)^{-\mathbf{1}}\boldsymbol{F} \qquad (26).$$

The corresponding model for the industry-by-industry dimension in found by transforming final demand for products from a specific region to demand for industry-output of that region by pre-multiplying with region specific sector-shares and adjusting the total requirement matrix:

$$\boldsymbol{g} \equiv \tilde{\tilde{\boldsymbol{D}}}(\tilde{\boldsymbol{U}}\boldsymbol{g} + \boldsymbol{F}) \rightarrow \boldsymbol{p} = \left(\boldsymbol{I} - \tilde{\tilde{\boldsymbol{D}}}\tilde{\boldsymbol{U}}\right)^{-1}\tilde{\tilde{\boldsymbol{D}}}\boldsymbol{F} \quad (27).$$

As all of the four accounting frameworks presented in this subsection are related to the same kind of input-output models, the question of which framework should be selected remains. According to Jackson and Schwarm [2011] the use-regionalized framework is argued to correspond more closely to a demand-driven model, whereas the make-regionalized framework corresponds to a supply-driven model, as it is implied that an increase in industry output induces an increase in demand for output of that industry in regions and purchasing industries. Furthermore, in the case of the German table data from income- and expenditure survey on households as well as from industrial cost structure surveys are available, so that for the use of these superior data sources a use-regionalized framework in more favorable. The German Multi-Regional Supply-Use Table (GerMRSUT) will, therefore, be based on the use-regionalized framework as presented in table 3.

Table 5Supply and Use-regionalized inter-regional framework

		Region r			Region s			Exports (row)	Total
		Product	**Industry**	**Final Demand**	**Product**	**Industry**	**Final Demand**		
Region r	**Product**		$\boldsymbol{U}^{rr}$	$\boldsymbol{d}^{rr}$		$\boldsymbol{U}^{rs}$	$\boldsymbol{d}^{rs}$	$\boldsymbol{e}^{r}$	$\boldsymbol{p}^{r} = \boldsymbol{s}^{r} - \boldsymbol{i}\boldsymbol{V}^{sr} - \boldsymbol{m}^{r}$
	Industry	$\boldsymbol{V}^{rr}$			$\boldsymbol{V}^{rs}$				$\boldsymbol{g}^{r}$
Region s	**Product**		$\boldsymbol{U}^{sr}$	$\boldsymbol{d}^{sr}$		$\boldsymbol{U}^{ss}$	$\boldsymbol{d}^{ss}$	$\boldsymbol{e}^{s}$	$\boldsymbol{p}^{s} = \boldsymbol{s}^{s} - \boldsymbol{i}\boldsymbol{V}^{rs} - \boldsymbol{m}^{s}$
	Industry	$\boldsymbol{V}^{sr}$			$\boldsymbol{V}^{ss}$				$\boldsymbol{g}^{s}$
Imports (row)		$\boldsymbol{m}^{r}$	$\boldsymbol{U}^{row\ r}$	$\boldsymbol{d}^{row\ r}$	$\boldsymbol{m}^{s}$	$\boldsymbol{U}^{row\ s}$	$\boldsymbol{d}^{row\ s}$		
Value Added			$\boldsymbol{w}^{r}$			$\boldsymbol{w}^{s}$			
Total		$\boldsymbol{s}^{r}$	$\boldsymbol{g}^{r}$		$\boldsymbol{s}^{s}$	$\boldsymbol{g}^{s}$			

Source: own calculations IEK-STE 2013

III The Non-Survey Multi-Regional Supply-Use Table for Germany

The estimation of the non-survey version of the German Multi-Regional Supply-Use Table (GerMRSUT) is performed in two broad steps. At first, 16 individual single regional supply-use tables for each federal state are constructed. These individual tables are estimated by means of regionalization of a national supply-use table for 2007. The estimation of the single regional tables can be distinguished into three major parts according to data sources and methods used: Firstly, the estimation of regional supply tables, value-added and the intermediate part of the use-tables is mostly based on partial information about regional employment, value-added and output. Secondly, as in contrast to the national table the household sector is separated into six income groups, estimation of final demand of these household sectors is based on the income and expenditure survey on households. Finally, for the estimation of a region's gross trade flows with the rest of the country and the rest of the world a modified version of CHARM [Kronenberg, 2009]approach is developed and applied.

III.1 Estimation of 16 Single-Regional Tables

In the case of the national supply-use table it was possible to choose between tables published by the federal statistical office and tables form the World Input-Output Database (WIOD). For this project it was decided to regionalize the WIOD supply-use tables for following reasons:

First, for the assessment of economic impacts due to political actions or infrastructure projects use tables should preferably be valued at basic prices instead of purchaser prices. In contrast to the latter basic prices do not include net taxes on products as well as trade margins[5], instead they are separated and reallocated to trade industries and a net taxes on products account by industry. Despite of being more closely related to economic theory of production valuation at basic prices has the advantage, that effects on net taxes on products or trade industries can be studied directly [Eurostat/European Commission, 2008]. The federal statistical office only publishes use tables at purchaser prices along with a tax and a trade margin vector, whereas WIOD provides use tables at basic prices as well as two margin tables.

Second, considering latter applications on regional and social distribution effects of energy policy, WIOD provides the more comprehensive data environment, as additional accounts for material use, energy consumption, skill-content of labour etc. are provided. Furthermore, for the national level time series of supply-use tables a current and previous year prices are available, which allow the estimation of price elasticities for the application of dynamic input-output or general equilibrium models.

[5] In the official German SUTs as well as in WIOD SUTs for Germany explicit information on transport margins is published.

III.1.1 Supply and Intermediate Use

For the estimation of regional industry output and value-added components following data were provided by the working group for national accounts of the federal states and the federal agency of employment:

- Value-added and wages were provided for all states at varying levels of detail ranging from 16NACE Rev. 1.1 Sections (A to P) to 59 2-digit categories.
- Gross output was provided for 16 NACE Rev. 1.1 Sections (A to P) for all states except for Mecklenburg West-Pomerania and Thuringia.
- Employment data was provided at a 3-digit level for all states.

The WIOD SUTs are based on SUTs and national accounts provided by statistical agencies of participating countries, but they've run through several procedures of harmonization, especially with respect to international trade relationships. Total value-added and gross output on the level of 16 broad NACE Rev. 1.1 industry sections are consistent with corresponding values from official data sources, but there are small deviations in intermediate and final consumption structures as well as in the distribution of gross output and value-added over industries on a 2-digit level of disaggregation.

In a first step regional value added data was split into individual industry-by-region tables one for each NACE Rev. 1.1 section that exhibits a higher level of disaggregation in the WIOD SUTs. This was the case for manufacturing (D), trade and repair (G), transport, storage and communication (I) and renting, real estate and business activities (K). For regions that provide value-added data at a 2-digit level, some information had to be aggregated to the WIOD level of disaggregation, whereas for regions that provide data only for NACE Rev. 1.1 sections missing entries have been estimated by assuming that value-added of an industry in proportional to employment as:

$$w_{i\cdot}^{r} = w_{i\cdot}^{n} * (empl_{i}^{r} / empl_{i}^{n}) \quad (28)$$

, where $w_{i\cdot}^{r}$ and $w_{i\cdot}^{n}$ denote value-added of industry i in region r or in the nation n and $empl_{i}^{r}$ and $empl_{i}^{n}$ denote employment of that industry on a regional or national level. As row and column sums are consistent neither with the national WIOD totals nor to regional totals in the case of estimated entries, RAS was applied for the adjustment to row and column totals. Regional wages were estimated on analogy to the procedure for value-added data. The results of this procedure are value-added and wage data that are consistent with national WIOD totals as well as with regional totals for NACE Rev. 1.1 Sections.

Other value-added than wages is calculated as the difference between value-added and wages by industry and further disaggregated into net operating surplus, depreciation and net taxes on production by assuming the regional shares of these components in other value-added by industry to be equal to the national shares taken from the official use table[6]:

[6] WIOD use tables only report wages and other value-added.

$$w_{il,(l\neq 1)}^{r} = w_{il,(l\neq 1)}^{n} * (w_{i.}^{r} - w_{i1}^{r})/(w_{i.}^{n} - w_{i1}^{n}) \quad (29).$$

The estimation of regional output by industry is carried out in two steps. For all regions despite Mecklenburg West-Pomerania and Thuringia information about output is available for NACE Rev. 1.1 sections, so that the first step consists of the estimation of output for the two missing regions. This is done by allocating the difference between national WIOD totals and the sum of output of the other states according to Mecklenburg West-Pomerania's and Thuringia's share of value-added for each section. As for the estimation of value added and wages individual industry-by-region tables for the sections D, G, I and K are constructed for the estimation of output by industry for 2-diggit NACE Rev. 1.1 categories. The entries of those tables are estimated by assuming that output by industry in a region is proportional to regional value-added of that industry, analogue to (28) for value-added:

$$g_{i\cdot}^{r} = g_{i\cdot}^{n} * (w_{i}^{r}/w_{i}^{n}) \quad (30)$$

This first approximation is then adjusted to national WIOD and to regional totals of a NACE section via RAS.

Given the estimates for regional output and value-added by industry it is possible to calculate intermediate consumption at purchaser prices as $l_{i,\ p.p.}^{r} = g_{i}^{r} - w_{i}^{r}$. For the estimation of total intermediate consumption at basic prices it is necessary to deduce net taxes on products. This is carried out by assuming that tax payments by industry are proportional to total intermediate consumption at purchaser prices, so that $tax_{i}^{r} = tax_{i}^{n}\left(l_{i,\ p.p.}^{r}/l_{i,\ p.p.}^{n}\right)$. Total intermediate consumption at basic prices is then calculated as $l_{i,\ b.p.}^{r} = l_{i,\ p.p.}^{r} - tax_{i}^{r}$.

The structure of intermediate consumption in region r is estimated by stating the assumption that each industry has the same composition of purchases on the regional and national level, so that:

$$u_{ji}^{r} = u_{ji}^{n}\left(l_{i,\ b.p.}^{r}/l_{i,\ b.p.}^{n}\right) \quad (31).$$

It is preferable to use regional shares of total intermediate consumption instead of shares in regional output or value-added, as the formulation in (31) better reflects regional differences in productivity, which are known as fabrication effects [Round, 1978].

Finally, for the estimation of regional production structures of industries for the supply table it was assumed that the composition of products produced by an industry is equal to the national level. The elements of the national supply table are therefore scaled by the share of industry i's output in region r in national output of that industry:

$$v_{ij}^{r} = v_{ij}^{n}(g_{i}^{r}/g_{i}^{n}) \quad (32).$$

After these steps regional supply matrices, as well as regional value added and intermediate consumption tables of the initial single regional tables are populated. The remaining items to be estimated are final demand and interregional as well as foreign trade.

III.1.2 Domestic Final Use

Domestic final demand is subdivided into consumption of households, non-profit organizations serving households (NPISH), and government, as well as gross capital formation. For the estimation of the consumption of regional governments and gross capital formation regional totals are available from regional accounts, but their national total deviates from those given by the WIOD use table. These items are first adjusted to match national totals of the WIOD use table and then applied to the national product structures by assuming equal structure on the national and the regional level. The adjustment is proceeded according to:

$$\hat{d}^{r}_{k\cdot\,p.p.} = d^{r}_{k\cdot\,p.p.}\left(d^{n\,WIOD}_{k\cdot\,p.p.}/d^{\cdot}_{k\cdot\,p.p.}\right) \quad (33),$$

where k denotes a category of final demand, $d^{n\,WIOD}_{k\cdot\,p.p.}$ denotes the national total at purchaser prices and $d^{\cdot}_{k\cdot\,p.p.}$ the corresponding sum of regional totals. For the consumption of non-profit organizations serving households regional shares in national population are used to estimate total regional consumption at purchaser prices of that category, as there is no information available. On analogy to the deduction of net taxes on products for total intermediate consumption, it is assumed that tax payments are proportional to total expenditures at purchaser prices: $tax^{r}_{k} = tax^{n}_{k}\left(\hat{d}^{r}_{k\cdot\;p.p.}/d^{n}_{k\cdot\;p.p.}\right)$. Afterwards, total consumption at basic prices is calculated as $\hat{d}^{r}_{k\cdot\,b.p.} = \hat{d}^{r}_{k\cdot\,p.p.} - tax^{r}_{k}$. In final step expenditures for individual product groups are estimated analogue to intermediate consumption by assuming that regional structures are equal to those of the national use-table:

$$d^{r}_{kj} = d^{n}_{kj}\left(\hat{d}^{r}_{k\cdot\,b.p.}/d^{n}_{k\cdot\,b.p.}\right) \quad (34).$$

Due to the crucial role of consumption of households and the later application of the German MRSUT to distributional effects of the German energy policy on disposable income and consumption of different income groups, consumption structures are estimated from the income and expenditure survey (EVS for Einkommens- und Verbrauchsstichprobe) for 2008. The EVS is conducted every five years and based on a sample of approx. 60,000 households. Participants allocate their gross income on six types namely wage income, profit income, rental income, returns from monetary assets, public and non-public transfers, whereas consumption expenditures are allocated on 133 types of intended use COICOP categories. In the regional supply-use tables final demand categories $k = 1, \dots, 6$ represent those income-groups. The estimation is carried out in three broad steps, which include, first, the estimation of consumption by product of regional households, second, the disaggregation of expenditures according to six types of income and, finally, the conversion of expenditures at purchaser prices into basic prices.

Expenditures of each household by COICOP categories are first weighted by an expansion factor, which indicates the number of households represented by a household participating in the survey. Thereafter consumption expenditures are aggregated according to regions. Because expenditure surveys are often affected incorrectly filled questionnaires [Lehmann, 2004], adjustments via correction coefficients for each COICOP category are calculated as quotients of expenditures for a COICOP category from national accounts and the corresponding national aggregates calculated from EVS data. These correction coefficients are, then, applied to regional values such that it is assumed that the degree deviation within each COICOP category is invariant with respect to regions. In order to integrate the consumption data into the regional use tables, it is necessary to convert the data from COICOP categories into CPA product groups via the most recent consumption interdependence table, which was publish for 2006 by the federal statistical office [Kronenberg & Többen, 2011]. This step delivers preliminary estimates of consumption expenditures for 2008 valuated at purchaser prices $\bar{d}^{r}_{j,\ p.p.}$.For the deflation to 2007 and for the correction of deviations due to the use of the 2006 consumption interdependence table preliminary estimates are arranged in a product-by-region table, so that row sums and column sum from the national consumption of product j and total consumption of households in region r respectively. The entries are then adjusted to the 2007 national consumption vector of households taken from the WIOD use table at purchaser prices and to regional totals of consumption of households taken from regional account data via RAS, such that $d^{n}_{j,\ p.p.} = \sum_r \bar{d}^{r}_{j\,p.p.}$ and $d^{r}_{.,\ p.p.} = \sum_j \bar{d}^{r}_{j\,p.p.}$.

The second step consists of disaggregating regional consumption expenditures by product into six types of income. As many households have several types of income, consumption expenditures for COICOP categories of each household are divided according to the shares of each income type in household's total income. These expenditures are, then, aggregated with respect to the type of income and location, which yields regional expenditures by product $\tilde{d}^{r}_{jk\,p.p.}$ for each region and income-group. Similar to the first step expenditures are converted from COICOP categories into CPA products, which are further used to calculate coefficients for the disaggregation $\theta^{r}_{jk} = \tilde{d}^{r}_{jk\,p.p.} / \sum_{k=1}^{6} \tilde{d}^{r}_{jk,\ p.p.}$ for $k = 1, \dots, 6$. Finally, the RAS adjusted expenditure are allocated on types of income according to:

$$d^{r}_{jk\,p.p.} = \theta^{r}_{jk} \bar{d}^{r}_{j\,p.p.} \tag{35}$$

This procedure is based on the assumption that consumption expenditures of an individual household only depend on total income. Note, that it also allows individual consumption structures of households for each region and type of income.

As final step a revaluation from purchaser to basic prices is conducted by stating the assumption that the share of net taxes and trade margins in household's consumption of product j at purchaser prices is equal among all regions and types of income. Information on net taxes and trade margins embodied in consumption expenditures of households by product are delivered by the WIOD valuation matrices. First, net taxes on products are deduced as:

$$d^{r}_{jk\,(p.p.-tax)} = d^{r}_{jk\,p.p.} \left(1 - \overline{tax}^{n}_{j} / d^{n}_{j,\ p.p.}\right) \tag{36}$$

Thereafter, trade margins are deduced from each type of product consumed and reallocated on the corresponding type of trade service, which yields expenditures by product at basic prices for each region and type of income:

$$d^{r}_{jk\,b.p.} = d^{r}_{jk\,(p.p.-tax)} \left(1 - \frac{margin^{n}_{j}}{d^{n}_{j,\ p.p.} - \overline{tax}^{n}_{j}}\right) \tag{37}$$

Thus, the procedure for the conversion of purchaser into basic prices, is based on the assumptions that net tax rates and the share of trade margins of a product is invariant with respect income type and region. After this step single regional use-tables can be populated with final consumption of households disaggregated into six types of income, so that only interregional and foreign trade remains to be estimated.

III.1.3 Regional Trade

Originally it was planned to base the estimation of regional foreign trade on regional trade statistics that report foreign imports and exports of the federal state and distinguishes 205 3-digit EWG 2002[7] product groups. The use of this information was hampered for several reasons: First, for each product group there is a share imports and exports that is not allocated to a specific region of origin or destination but to "the rest", as there is not enough reliable information about region of consumption or production. For some products over 90% of trade is allocated to the rest. Second, especially for regions, which are a central trading hub such as Hamburg and Bremen, imports and exports are not consistent with data on regional production and consumption. For these reasons regional trade statistics are used for the initial estimate, but for the construction of constraints attached with high standard deviations, in order to reflect the poor data quality. Instead, data on national foreign trade embodied in the WIOD SUTs is allocated to the regions for the initial estimate.

The first step of this procedure consists of eliminating re-exports from the national WIOD foreign trade vectors. Comparing national foreign imports and exports reported in the WIOD SUTs with those form the official tables, reveals only small deviations due to harmonization adjustments. The official import table reports a certain amount of re-exports for all physical products and some services[8], whereas the WIOD import table contains only zeros. Furthermore, for some products such as pharmaceuticals or office machinery imports exceed total domestic consumption and exports domestic output. For the elimination of re-exports shares of re-exports in total exports are calculated from the official import table and applied the foreign trade columns of the WIOD SUTs.

As a second step regional output of product j p_j^r is interpreted as export potential, whereas total regional domestic use of product j $u_{j.}^r + d_{j.}^r$is interpreted as import potential. Regional trade with the rest of the world is then estimated as:

$$e_j^{r\ row} = e_j^{r\ row} \frac{p_j^r}{p_j^n} \tag{38}$$

---

[7] EWG is a German product classification for foreign trade and is not derived from any UN classification, as its first version was established 1936. However, the conversion into 2-digit CPA product groups is possible without conflicts.

[8]for some products such as wearing apparels or office machinery more than 60% of exports are consist of imported products

$$m_j^{row\,r} = m_j^{row\,n} \frac{u_{\cdot j}^r + d_{\cdot j}^r}{u_{\cdot j}^n + d_{\cdot j}^n} \tag{39}$$

Regional gross imports and exports are estimated via a refined version of Kronenberg's [Kronenberg, 2009] CHARM procedure. The refinement is necessary due to the fact that the original CHARM procedure may yield estimates for regional imports and/or exports, which exceed total regional output or domestic consumption. In the original CHARM formula the parameter h_j plays a crucial role in the estimation procedure and is calculated from observed national trade flows with the rest of the world:

$$h_j^n = \frac{vol_j^n - |b_j^n|}{\left(p_j^n + u_{\cdot j}^n + d_{\cdot j}^n\right)} = \frac{q_j^n}{\left(p_j^n + u_{\cdot j}^n + d_{\cdot j}^n\right)} \tag{40}$$

, where $q_j^n \equiv vol_j^n - |b_j^n|$ denotes the amount of cross-hauling, which is defined as the difference of the trade volume of the nation with the rest of the world $vol_j^n = e_j^{n\,row} + m_j^{row\,n}$and the absolute value of the national trade balance$|b_j^n| = |e_j^{n\,row} - m_j^{row\,n}|$. h_j^n is, therefore, the share of cross-hauling in the sum of total domestic output and consumption, which is then interpreted as a proxy for the heterogeneity of product j. Kronenberg [2009] assumes that product heterogeneity is characteristic to a product and not to geographical location, so that $h_j^n = h_j^r$ may be applied to the sum of regional domestic output and consumption, in order to estimate regional cross-hauling:

$$q_j^r = h_j^r\left(p_j^r + u_{\cdot j}^r + d_{\cdot j}^r\right) \tag{41}$$

This formula implies that the share cross-hauling in the sum of domestic output and consumption in the region is equal to that of the nation, whereby q_j^r consists of regional trade with the rest of the country as well as with the rest of the world. Regional gross imports and exports may then be calculated as:

$$e_j^{r\cdot} = \frac{vol_j^r + b_j^r}{2} = \frac{q_j^r + |b_j^r| + b_j^r}{2} \tag{42a}$$

$$m_j^{r\cdot} = \frac{vol_j^r - b_j^r}{2} = \frac{q_j^r + |b_j^r| - b_j^r}{2} \tag{42b}.$$

As afore mentioned does this procedure yield estimates for imports and exports that exceed regional output ($p_j^r < e_j^{r\cdot}$) and consumption ($u_{\cdot j}^r + d_{\cdot j}^r < m_j^r$)[9].This effect can occur for several reasons, which result from the estimation formula for h_j^n:

- If h_j^n is calculated from national trade data which embody re-exports to such an extent that national cross-hauling is greater than the sum of output and domestic consumption, $h_j^n > 1$ and resultsin regional cross-hauling that exceeds the sum of re-

[9] Note that as $b_j \equiv e_j - m_j \equiv p_j - u_{\cdot j} - d_{\cdot j}$implies $p_j - e_j = u_{\cdot j} + d_{\cdot j} - m_j$ the cases of $p_j^r < e_j^{r\cdot}$ and $u_{\cdot j}^r + d_{\cdot j}^r < m_j^r$ must occur simultaneously.

gional output and domestic consumption. For this reason re-exports should be eliminated from national trade for the application of CHARM to ensure $h_j^n \leq 1$.

- If a product is either not produced or consumed in a region ($p_j^r = 0$ or $u_{\cdot j}^r + d_{\cdot j}^r = 0$), but cross-hauling is observed in national foreign trade ($h_j^n > 0$), CHARM will yield $e_j^{r\cdot} > 0$ in spite of $p_j^r = 0$ or $m_j^r > 0$ in spite of $u_{\cdot j}^r + d_{\cdot j}^r = 0$. Therefore,itis required that$h_j^r = 0$for $min(p_j^r; u_{\cdot j}^r + d_{\cdot j}^r) = 0$.

These two cases are the more obvious ones for which the problem of $p_j^r < e_j^{r\cdot}$ and $u_{\cdot j}^r + d_{\cdot j}^r < m_j^r$ occurs, but such inconsistencies are still possible even if re-exports are eliminated from national trade data andh_j^r is set to zero for $min(p_j^r; u_{\cdot j}^r + d_{\cdot j}^r) = 0$. This problem arises from the requirements 3 and 4 stated by Kronenberg [2009 for the justification of equation (41). These requirements claim that an increase in either domestic output or domestic consumption while consumption or output remains constant will lead to an increase in regional exports or imports respectively. But an increase in regional output also increases intermediate consumption, which is likely to be satisfied to a certain degree by additional imports. A similar argument is made for an increase in domestic consumption.

Although these arguments are plausible, formula (41) can yield inconsistencies, because of the linkage between domestic output and regional imports and domestic consumption and regional exports. In order to verify this argument formally it is necessary to distinguish the cases $b_j^r < 0$and $b_j^r > 0$:

- If $b_j^r < 0$ (42a) becomes $e_j^{r\cdot} = \frac{q_j^r}{2}$. In order to show in which cases exports exceed output, (41) is substituted into the relation $e_j^{r\cdot} = \frac{q_j^r}{2} > p_j^r$ and further solved for h_j^ryielding:

$$h_j^r > \frac{2p_j^r}{(p_j^r + u_{\cdot j}^r + d_{\cdot j}^r)} \qquad (43a).$$

- On analogy to the derivation of (42a) in the case of $b_j^r > 0$, (41) is substituted into the relation $m_j^{r\cdot} = \frac{q_j^r}{2} > u_{\cdot j}^r + d_{\cdot j}^r$ and solved for h_j^r, which yields:

$$h_j^r > \frac{2(u_{\cdot j}^r + d_{\cdot j}^r)}{(p_j^r + u_{\cdot j}^r + d_{\cdot j}^r)} \qquad (43b).$$

The fact that $h_j^n = h_j^r$ is assumed allows to substitute (41) into (43a) and (43b), which leads to:

$$\frac{q_j^n}{(p_j^n + u_{\cdot j}^n + d_{\cdot j}^n)} > \frac{2p_j^r}{(p_j^r + u_{\cdot j}^r + d_{\cdot j}^r)} <=> \frac{q_j^n}{2} \frac{(p_j^r + u_{\cdot j}^r + d_{\cdot j}^r)}{(p_j^n + u_{\cdot j}^n + d_{\cdot j}^n)} > p_j^r \qquad (44a)$$

$$\frac{q_j^n}{(p_j^n + u_{\cdot j}^n + d_{\cdot j}^n)} > \frac{2(u_{\cdot j}^r + d_{\cdot j}^r)}{(p_j^r + u_{\cdot j}^r + d_{\cdot j}^r)} <=> \frac{q_j^n}{2} \frac{(p_j^r + u_{\cdot j}^r + d_{\cdot j}^r)}{(p_j^n + u_{\cdot j}^n + d_{\cdot j}^n)} > (u_{\cdot j}^r + d_{\cdot j}^r) \qquad (44b).$$

As the relationship of $q_j^n/2$ to national imports and exports depends on the sign of the national trade balance, it is necessary to further distinguish negative and positive trade balances for (44a) and (44b) each. Because $b_j^n > 0$ implies $q_j^n/2 = m_j^n$ and for $b_j^n < 0$ $q_j^n/2 = e_j^n$, following cases yielding regional imports and exports that exceed regional output and consumption can, finally, be derived:

$$e_j^n \frac{\left(p_j^r + u_{\cdot j}^r + d_{\cdot j}^r\right)}{\left(p_j^n + u_{\cdot j}^n + d_{\cdot j}^n\right)} > p_j^r \qquad \text{for } b_j^n > 0 \text{ and } b_j^r < 0 \qquad (45a')$$

$$m_j^n \frac{\left(p_j^r + u_{\cdot j}^r + d_{\cdot j}^r\right)}{\left(p_j^n + u_{\cdot j}^n + d_{\cdot j}^n\right)} > p_j^r \qquad \text{for } b_j^n < 0 \text{ and } b_j^r < 0 \qquad (45a'')$$

$$e_j^n \frac{\left(p_j^r + u_{\cdot j}^r + d_{\cdot j}^r\right)}{\left(p_j^n + u_{\cdot j}^n + d_{\cdot j}^n\right)} > \left(u_{\cdot j}^r + d_{\cdot j}^r\right) \qquad \text{for } b_j^n > 0 \text{ and } b_j^r > 0 \qquad (45b')$$

$$m_j^n \frac{\left(p_j^r + u_{\cdot j}^r + d_{\cdot j}^r\right)}{\left(p_j^n + u_{\cdot j}^n + d_{\cdot j}^n\right)} > \left(u_{\cdot j}^r + d_{\cdot j}^r\right) \qquad \text{for } b_j^n < 0 \text{ and } b_j^r > 0 \qquad (45b'').$$

Given these six cases it remains to clarify why the occurrence of $p_j^r < e_j^{r\cdot}$ and $u_{\cdot j}^r + d_{\cdot j}^r < m_j^r$ is a serious problem both from a theoretical as well as from a practical point of view:

The theoretical problem is mostly related to the estimation of h_j^n from national trade data that contain re-exports and its interpretation as a proxy for product heterogeneity. For this argument it is important to highlight the conceptual difference of simultaneous importation and exportation with imported and domestically produced products. According to the "International Merchandise Trade Statistics: Supplement to the Compilers Manual" from Statistic Division of the UN [2008] it is crucial to re-exports that products are not substantially transformed and it is, therefore, to be assumed that products involved in such trade are almost perfectly homogeneous. In contrast to this cross-hauling with domestically produced products is seen as the result of product differentiation and brand preferences (see Leigh [1970, Isserman [1980 and Norcliffe [1983). Applying (40) to national gross trade flows that contain re-exports is, thus, a contradiction to the interpretation of h_j^n as a proxy for product heterogeneity.

In view that (multi)regional SUTs usually build the base for regional Input-Output models the occurrence of $p_j^r < e_j^{r\cdot}$ and $u_{\cdot j}^{\cdot r} + d_{\cdot j}^{\cdot r} < m_j^{\cdot r}$ cannot simply be ignored. When it comes for example to the estimation of intraregional purchases on the basis of such regional trade estimates the resulting intermediate and final consumption of domestically produced products will be negative, if re-exports are ignored. Alternatively regional purchase coefficients, which account for re-exports in regional trade explicitly [Lahr, 2001], can be used to derive intraregional purchases, but this would require the estimation of regional re-exports. For the case that regional tables are constructed for a multi-regional SUT, as it is the case here, the task becomes even more complicated, because it is necessary to allocate re-exports to their regional origin of production and destination of consumption. Furthermore, it is necessary to split regional trade flows into trade with foreign countries and with the rest of the country under study. Accounting for re-exports would, in the end, complicate the whole construction of (multi)regional SUTs unnecessarily and add only little explanatory power to the final table.

A refined version of CHARM is, therefore, developed, which yields regional export and import estimates consistent with the assumption of zero re-exports. For the refinement it is important to note, that there is an upper limit for each regionally produced product, to which cross-hauling caused by product differentiation is possible. Each trade flow exceeding $min\left(p_j^r; u_{\cdot j}^r + d_{\cdot j}^r\right)$ cannot be explained by product heterogeneity, so that the upper limit for regional exports or imports involved in regional cross-hauling is:

$$\min(p_j^r; u_{\cdot j}^r + d_{\cdot j}^r) = \max(\min\left(e_j^{r\cdot}; m_j^{\cdot r}\right)) \quad (46).$$

Equation (46) can be interpreted as the maximum cross-hauling potential of a region or a nation. Contrary to the original CHARM formula, where h_j^n is calculated as the share of national cross-hauling in the sum of output and domestic consumption, $\tilde{h}_j^n$ in the modified formula is calculated as the share of cross-hauling observed in international trade in the maximum cross-hauling potential:

$$\tilde{h}_j^n = \frac{q_j^n}{2\min(p_j^r; u_{\cdot j}^r + d_{\cdot j}^r)} \quad (47)$$

The denominator is multiplied by two as q_j^n consists of both, exports and imports. If one assumes $\tilde{h}_j^n = \tilde{h}_j^r$, i.e. that the share of regional cross-hauling in the regional cross-hauling potential is the same as in the nation, regional cross-hauling may be estimated as:

$$q_j^r = 2\tilde{h}_j^r * \min(p_j^r; u_{\cdot j}^r + d_{\cdot j}^r) \quad (48)$$

Regional gross imports and exports are then calculated according to (42a) and (42b).

If, in addition, data or estimates on regional trade with the rest of world are available the splitting of trade with the rest of the world and the rest of the country can be done within the CHARM procedure. For the integrated estimation of regional trade with the rest of the world and the rest of the nation it is important, that regional foreign trade does not contain any re-exports. This approach is used for the initial estimate of the German MRSUT.

In the first step, the share of cross-hauling of regional foreign trade in the regional cross-hauling potential is calculated instead of the national counterpart as in (46):

$$\tilde{h}_j^{r\,row} = \frac{q_j^{r\,row}}{2\min(p_j^r; u_{\cdot j}^r + d_{\cdot j}^r)} \quad (49).$$

It is, then, assumed that $\tilde{h}_j^{r\,row} = \tilde{h}_j^{r\,roc}$, which seems to be a more plausible assumption, as product heterogeneity is rather a characteristic of diversification of regional industries than of the product group itself.$\tilde{h}_j^{r\,roc}$can, thus, be expected to capture special features of regional industries and thereby the heterogeneity of regional product output more properly. In the second step for the estimation of regional cross-hauling in trade flows with the rest of the country, $\tilde{h}_j^{r\,roc}$ is applied to the remaining regional cross-hauling potential after accounting for regional foreign trade:

$$q_j^{r\,roc} = 2\tilde{h}_j^{r\,roc} * \min(p_j^r - e_j^{r\,row}; u_{\cdot j}^r + d_{\cdot j}^r - m_j^{row\,r}) \quad (50).$$

The subtraction of regional foreign trade in (50) ensures that the estimate is consistent with $p_j^r < e_j^{r\cdot}$ and $u_{\cdot j}^{\cdot r} + d_{\cdot j}^{\cdot r} < m_j^{\cdot r}$. Regional gross imports and exports are then calculated analogously to (42a) and (42b) as:

$$e_j^{r\,roc} = \frac{q_j^{r\,roc} + \left|b_j^{r\,roc}\right| + b_j^{r\,roc}}{2} \quad (51a)$$

$$m_j^{roc\,r} = \frac{q_j^{r\,roc} + \left|b_j^{r\,roc}\right| - b_j^{r\,roc}}{2} \quad (51b),$$

Whereby $b_j^{r\,roc} = p_j^r + m_j^{row\,r} - u_{\cdot j}^r - d_{\cdot j}^r - e_j^{r\,row}$ denotes the interregional trade-balance of region r. With the estimation of interregional gross trade the construction of the single regional tables for the German is finished. The next step consists of the allocation of regional gross trade estimates to the origin of production and the destination of consumption.

III.2 Assembling of the Multi-Regional Non-Survey Table

The prior steps of the construction of the initial estimate were aiming at the estimation of 16 individual SUTs, one for each German state. The result of these steps is presented in table 2. The two final steps of the construction process consist of (1.) linking the 16 single-regional SUTs with respect to the interregional trade flows among the states and (2.) the distribution on domestically produced and imported products on industries and final demand sectors.

For the further allocation of regional gross exports to their regional destination and of gross imports to their regional origin, it is crucial to understand what kind of information CHARM delivers. That is, for each product CHARM delivers row and column sums of an origin-destination (OD) matrix, whose diagonal elements are zero, so that the task of the first step is to estimate the off-diagonal elements of those matrices. Such an OD matrix is presented in table 3. Note that, as the data on which CHARM is based are ether national totals of shares in national totals, regional gross trade estimates are consistent in the sense, that the sum of regional exports to the rest of the country equals the sum of regional imports from the rest of the country for each product j: $\sum_r e_j^{r\,roc} = \sum_r m_j^{roc\,r}\ \forall j$.

For the estimation of the off-diagonal elements it is assumed that the geographical distance between two states has no influence on scale of bilateral trade flows. For the final MRSUT the effect of distance is captured by the constraints on interregional trade flows, which are based on transportation data. As a first step the regional origin of imports from the rest of the country are distributed to the regions of origin according to their market-share in total interregional exports except exports of the importing state:

$$\tilde{t}_j^{rs} = \frac{m_j^{roc\,s}}{\sum_{r \neq s} e_j^{r\,roc}} \quad (52)$$

In final step of the construction procedure of the initial estimate total purchases of industries $u_{ji}^{\cdot r}$and domestic final demand sectors $d_{jk}^{\cdot r}$ are split with according to their geographical origin. For this purpose regional purchase coefficients are calculated for intraregional purchases, purchases from other regions and from the rest of the world. Due to the estimation procedure of regional with the rest of the country and the rest of the world, the existence of re-exports can be ruled out, so that according to [Lahr, 2001]for region $r \neq s$:

$$rpc_j^{rr} = \frac{\left(p_j^r - \tilde{t}_j^{r\cdot}\right)}{\left(p_j^r + \tilde{t}_j^{\cdot r} + m_j^{row\,r} - \tilde{t}_j^{r\cdot} - e_j^{r\,row}\right)} \quad (53a)$$

$$rpc_j^{sr} = \frac{\tilde{t}_j^{sr}}{\left(p_j^r + \tilde{t}_j^{\cdot r} + m_j^{row\,r} - \tilde{t}_j^{r\cdot} - e_j^{r\,row}\right)} \quad (53b)$$

$$rpc_j^{row\,r} = \frac{m_j^{row\,r}}{\left(p_j^r + \tilde{t}_j^{\cdot r} + m_j^{row\,r} - \tilde{t}_j^{r\cdot} - e_j^{r\,row}\right)} \quad (53c).$$

As there is no further information about the propensity to consume imported products of individual industries or final demand sectors, it is assumed that for each product j in region r industries and domestic final demand sector have the same average import propensity. Due to this assumption it is possible to apply the regional purchase coefficients to total use of product j in region r by industry i or final demand sector k. For intraregional purchases $u_{ji}^{rr} = rpc_j^{rr} * u_{ji}^{\cdot r}$and $d_{jk}^{rr} = rpc_j^{rr} * d_{jk}^{\cdot r}$, for interregional purchases $(r \neq s) u_{ji}^{sr} = rpc_j^{sr} * u_{ji}^{\cdot r}$ and $d_{jk}^{sr} = rpc_j^{sr} * d_{jk}^{\cdot r}$ as well as for purchases from foreign countries $u_{ji}^{row\,r} = rpc_j^{row\,r} * u_{ji}^{\cdot r}$ and $d_{jk}^{row\,r} = rpc_j^{row\,r} * d_{jk}^{\cdot r}$.

Table 5 Origin-Destination Matrix with CHARM estimated row and column sums

	Region 1	Region 2	⋯	Region r	Total
Region 1	0	$\tilde{t}_j^{12}$	⋯	$\tilde{t}_j^{1r}$	$e_j^{1\,roc}$
Region 2	$\tilde{t}_j^{21}$	0	⋯	$\tilde{t}_j^{2r}$	$e_j^{2\,roc}$
⋮	⋮	⋮	⋱	⋮	⋮
Region r	$\tilde{t}_j^{r1}$	$\tilde{t}_j^{r2}$	⋯	0	$e_j^{r\,roc}$
Total	$m_j^{roc\,1}$	$m_j^{roc\,2}$	⋯	$m_j^{roc\,r}$	$\sum_r e_j^{r\,roc} = \sum_r m_j^{roc\,r}$

Source: own calculations IEK-STE 2013

This procedure generates a non-survey use regionalized multiregional supply-use table, but so far the table is not balanced. As the application of (52) does not insure the sum of exports to other regions to be equal to CHARM estimates of gross exports to the rest of the country, regional product-output is not equal to total use that product. For this reason matrix balancing methods such as RAS have to be applied.

IV Table Balancing and the Introduction of Superior Data

The non-survey initial estimate of GerMRSUT isn't balanced so far, as regional product-output is not equal to total use of that product. Furthermore, the non-survey table –despite consumption of households- is only based on very basic data from national and regional accounts. Also, the spatial dimension is still absent from the table, as regional trade relationships are not affected by the distance between two trading regions. For this reason the quality of the table should be improved by adjusting it to information given by superior data sources on the one hand. On the other hand the resulting table must fulfil the fundamental accounting balances of the use-regionalized framework and be consistent with data used for the non-survey procedure.

A balancing method that fulfils all of these requirements is the most recent RAS variant KRAS, which was developed by [Lenzen et al., 2009] and is the core of AISHA, a tool for the construction of (series) of multi-regional supply use tables [Geschke et al., 2011]. The final version of GerMRSUT should fulfil following constraints, defined by available data and accounting balances:

- Accounting balances of the use-regionalized framework must be fulfilled, meaning that regional product output must be equal to total use of that product and output by industry must be equal to total consumption of intermediate and primary inputs.
- If regional supply and interregional use tables are summed up over all regions the result must be equal to the national supply and use tables
- For each region output, value-added and wages by industry as well as total final consumption of regional households, regional governments and total net taxes on products must be equal to data from regional accounts.
- Regional imports from and exports to the rest of the world for physical goods must be adjusted to information from regional foreign trade statistics.
- The spatial structure of interregional trade in physical goods must be adjusted to transport flows estimated from an entropy maximizing model.
- Regional cost structures of manufacturing industries must be adjusted to information gained from the industrial cost structure survey.

AISHA enables to handle all of these constraints in an integrated approach. The next subsections will first describe the general RAS approach and its recent variant KRAS. Thereafter, processing of superior data is described followed by the implementation of balancing and adjustment of the table in AISHA.

IV.1 The RAS Methodology and it's recent Variant: KRAS

A crucial requirement of every SUT is that economic accounting balances are represented in its structure, such that row and column sums for products and industries must be equal, i.e. in the case of the German MRSUT balances of the use-regionalized framework must be met.

The non-survey initial estimate does so far violate these balancing equations, so that a method for matrix balancing has to be applied. One of the most basic and still most popular methods is the so called RAS technique, whereby $\boldsymbol{A}$ represents a matrix (e.g. of interindustry transactions or input coefficients) and $\boldsymbol{R}$ and $\boldsymbol{S}$ adjustment vectors that are calculated iteratively. RAS can be used for reconciling a non-survey input-output table with known row and column sums (i.e. with balancing conditions such as (1) and (2) for the single region case) or for updating input-output tables, when only partial information about row and column sums are given for a more recent year. The goal is, thus, to estimate a matrix $\boldsymbol{A}(\mathbf{1})$ that is consistent with given row and column sums, $\boldsymbol{u}(\mathbf{1}) = \boldsymbol{A}(\mathbf{1})\boldsymbol{i}$ and $\boldsymbol{v}(\mathbf{1}) = \boldsymbol{i}'\boldsymbol{A}(\mathbf{1})$ with $\boldsymbol{iv}(\mathbf{1}) = \boldsymbol{iu}(\mathbf{1})$, and deviates only to a minimal degree from the initial matrix $\boldsymbol{A}(\mathbf{0})$[Miller & Blair, 2009]. As afore mentioned, the final matrix $\boldsymbol{A}(\mathbf{1})$ in terms of an iterative calculation procedure, assuming $\boldsymbol{u}(\mathbf{1}) \neq \boldsymbol{u}(\mathbf{0})$ and $\boldsymbol{v}(\mathbf{1}) \neq \boldsymbol{v}(\mathbf{0})$. The iterative process starts with step $k = 0$ and the initial matrix $\boldsymbol{A}(\mathbf{0}) = \boldsymbol{A}$ and is then carried out in three successive steps:

1. $\boldsymbol{A}^{\boldsymbol{k}+\mathbf{1}} = \hat{\boldsymbol{r}}^{\boldsymbol{k}+\mathbf{1}}\boldsymbol{A}^{\boldsymbol{k}}$ with $\hat{\boldsymbol{r}}^{\boldsymbol{k}+\mathbf{1}} = \hat{\boldsymbol{u}}(\mathbf{1})\hat{\boldsymbol{u}}^{\boldsymbol{k}^{-\mathbf{1}}}$
2. $\boldsymbol{A}^{\boldsymbol{k}+\mathbf{2}} = \boldsymbol{A}^{\boldsymbol{k}+\mathbf{1}}\hat{\boldsymbol{s}}^{\boldsymbol{k}+\mathbf{1}}$ with $\hat{\boldsymbol{s}}^{\boldsymbol{k}+\mathbf{1}} = \hat{\boldsymbol{v}}(\mathbf{1})\hat{\boldsymbol{v}}^{\boldsymbol{k}+\mathbf{1}^{-\mathbf{1}}}$
3. Set $k = k + 1$ and return to step 1.

The first two steps may be summarized to $\boldsymbol{A}^{\boldsymbol{k}+\mathbf{2}} = \hat{\boldsymbol{r}}^{\boldsymbol{k}+\mathbf{1}}\boldsymbol{A}^{\boldsymbol{k}}\hat{\boldsymbol{s}}^{\boldsymbol{k}+\mathbf{1}}$. These three steps are successively repeated until deviations of row and column sums ($\boldsymbol{u}^{\boldsymbol{k}}and\boldsymbol{v}^{\boldsymbol{k}}$) from those known for the final matrix ($\boldsymbol{u}(\mathbf{1})$and $\boldsymbol{v}(\mathbf{1})$) are sufficiently small, this is: $\left|\boldsymbol{u}(\mathbf{1}) - \boldsymbol{u}^{\boldsymbol{k}}\right| \leq \varepsilon$ and $\left|\boldsymbol{v}(\mathbf{1}) - \boldsymbol{v}^{\boldsymbol{k}}\right| \leq \varepsilon$, where ε is an arbitrary small positive number. Generally for an infinite number of steps $\lim_{n\to\infty} \hat{\boldsymbol{r}}^{\boldsymbol{k}+\boldsymbol{n}}\boldsymbol{A}(\mathbf{0})\hat{\boldsymbol{s}}^{\boldsymbol{k}+\boldsymbol{n}} = \boldsymbol{A}(\mathbf{1})$, whereby $\hat{\boldsymbol{r}}^{\boldsymbol{k}+\boldsymbol{n}} = \prod_{i=1}^{n} \hat{\boldsymbol{r}}^{\boldsymbol{k}+\boldsymbol{i}}$ and$\hat{\boldsymbol{s}}^{\boldsymbol{k}+\boldsymbol{n}} = \prod_{i=1}^{n} \hat{\boldsymbol{s}}^{\boldsymbol{k}+\boldsymbol{i}}$ [Geschke et al., 2011]. Furthermore, the RAS procedure has two important properties: Firstly, signs of the initial matrix $\boldsymbol{A}(\mathbf{0})$are preserved, so that a final matrix estimated from a non-negative initial one will non-negative two. Second, zero elements of the initial matrix will remain zero in the final table [Miller & Blair, 2009].

The iterative RAS procedure can be interpreted as the solver of a constraint optimization problem [Bacharach, 1970]. Thereby, the objective consists of minimizing the loss of information given by the initial matrix, i.e. the differences between the initial $\boldsymbol{A}(\mathbf{0})$ and final table elements $\boldsymbol{A}(\mathbf{1})$, constrained by known row and column sums of the final matrix:

$$\min_{a_{ij}(1)} f\left(a_{ij}(0), a_{ij}(1)\right) = \sum_i \sum_j \left(a_{ij}(1) \ln\left[\frac{a_{ij}(1)}{a_{ij}(0)}\right]\right) \quad (54a)$$

s.t.

$$u_i(1) = \sum_j a_{ij}(1) \quad (54b)$$

$$v_j(1) = \sum_i a_{ij}(1) \quad (54c)$$

The corresponding Lagrange function appears as:

$$L = \sum_i \sum_j \left(a_{ij}(1) \ln\left[\frac{a_{ij}(1)}{a_{ij}(0)}\right]\right) + \sum_i \lambda_i\left[u_i(1) - \sum_j a_{ij}(1)\right] + \sum_j \mu_i\left[v_j(1) - \sum_i a_{ij}(1)\right] \quad (55),$$

where λ_i and μ_i are Lagrangian multipliers. The solution of the minimal information loss problem is then found by working out the related first-order conditions and solving the resulting system of equations:

$$\frac{\partial L}{\partial a_{ij}(1)} = 1 + \ln a_{ij}(1) - \ln a_{ij}(0) - \lambda_i - \mu_i = 0 \quad (56a)$$

$$\frac{\partial L}{\partial \lambda_i} = u_i(1) - \sum_j a_{ij}(1) = 0 \quad (57b)$$

$$\frac{\partial L}{\partial \mu_i} = v_j(1) - \sum_i a_{ij}(1) = 0 \quad (57c)$$

Solving (57b) for $a_{ij}(1)$ and rearranging yields:

$$a_{ij}(1) = e^{(\lambda_i - 1/2)} a_{ij}(0) e^{(\mu_i - 1/2)} \quad (58)$$

Defining $r_i = e^{(\lambda_i - 1/2)}$ and $s_j = e^{(\mu_i - 1/2)}$ delivers then the familiar relationship between the elements of the final and the initial matrix, with it's adjustment terms r_i and s_j:

$$a_{ij}(1) = r_i a_{ij}(0) s_j \quad (59).$$

Finally, the values for r_i and s_j are found by inserting (3) the two remaining first-order conditions (1b) and (1c):

$$r_i = \frac{u_i(1)}{\sum_j a_{ij}(0) s_j} \quad (60a)$$

$$s_j = \frac{v_j(1)}{\sum_i a_{ij}(0) r_i} \quad (60b)$$

Solving both equations iteratively delivers then the adjustment terms with which $\boldsymbol{A}(\boldsymbol{0})$ is to be multiplied, in order to reach the final table [Miller & Blair, 2009].

The application of the original RAS, however, is hampered by several limitations, which gave rise to numerous extensions of the original RAS procedure:

First, the conventional RAS is not able to handle negative entries, which arise frequently in input-output and supply-use tables such as changes inventories or net taxes on products. In this case the solution won't converge towards the optimal solution. This problem was solved by [Junius & Oosterhaven, 2003], who developed the so called GRAS algorithm (G stands for generalized), which is capable to deal with negative entries as well as with negative row and column sums. They solve the problem of dealing with negatives by restating the objective of minimal information loss:

$\tilde{a}_{ij} = \frac{a_{ij}(1)}{a_{ij}(0)}$ with $\tilde{a}_{ij} > 0$ for $a_{ij}(0) \neq 0$ and $\tilde{a}_{ij} = 0$ else, so that $a_{ij}(1) = |a_{ij}(0)| \tilde{a}_{ij}$

$$\min_{a_{ij}(1)} f\big(a_{ij}(0), a_{ij}(1)\big) = \sum_i \sum_j \big(|a_{ij}(0)| \tilde{a}_{ij} \ln[\tilde{a}_{ij}]\big) \quad (61a)$$

s.t.

$$u_i(1) = \sum_j |a_{ij}(0)| \tilde{a}_{ij} \quad (61b)$$

$$v_j(1) = \sum_i |a_{ij}(0)| \tilde{a}_{ij} \quad (61c)$$

The basic idea of GRAS becomes apparent in the Lagrangian function related to the restated optimization problem, in which the initial matrix elements $a_{ij}(0)$ are decomposed into a semi-positive and a negative subset:

$$L = \sum_{(i,j)\in P}\left(\left|a_{ij}(0)\right|\tilde{a}_{ij}\ln\left[\tilde{a}_{ij}\right]\right) - \sum_{(i,j)\in N}\left(\left|a_{ij}(0)\right|\tilde{a}_{ij}\ln\left[\tilde{a}_{ij}\right]\right) + \sum_i \lambda_i\left[u_i(1) - \sum_j a_{ij}(0)\tilde{a}_{ij}\right] + \sum_j \mu_i\left[v_j(1) - \sum_i a_{ij}(0)\tilde{a}_{ij}\right] \quad (62),$$

where P and N are the sets (i,j) for which $a_{ij}(0) \geq 0$ and $a_{ij}(0) < 0$, respectively. On analogy to the conventional RAS problem the solution is found by setting up the corresponding first-order condition with respect $\tilde{a}_{ij}$, so that the final matrix is obtained as:

$a_{ij}(1) = \left|a_{ij}(0)\right|\tilde{a}_{ij} = r_i a_{ij}(0)\, s_j/e$ for $a_{ij}(0) \geq 0$

$a_{ij}(1) = \left|a_{ij}(0)\right|\tilde{a}_{ij} = r_i^{-1} a_{ij}(0)\, s_j^{-1}/e$ for $a_{ij}(0) < 0$.

The adjustment terms r_i and s_j can, then, be calculated iteratively from the system of non-linear equations, that results from the first-order conditions of the Lagrangian multipliers λ_i and μ_i after inserting the solutions for $a_{ij}(1)$:

$$\sum_j\left[r_i a_{(i,j)\in P}(0)s_j - r_i^{-1}a_{(i,j)\in N}(0)s_j^{-1}\right] = e u_i(1) \quad (63a)$$

$$\sum_i\left[r_i a_{(i,j)\in P}(0)s_j - r_i^{-1}a_{(i,j)\in N}(0)s_j^{-1}\right] = e v_j(1) \quad (63b)$$

The iterative calculation process is terminated if $\left|(\hat{\boldsymbol{r}}\boldsymbol{A}(\boldsymbol{0})_{(i,j)\in P}\hat{\boldsymbol{s}})\boldsymbol{i} - \left(\hat{\boldsymbol{r}}^{-1}\boldsymbol{A}(\boldsymbol{0})_{(i,j)\in N}\hat{\boldsymbol{s}}^{-1}\right)\boldsymbol{i} - \boldsymbol{u}(\boldsymbol{1})\right| < \varepsilon|\boldsymbol{u}(\boldsymbol{1})|$ and $\left|(\hat{\boldsymbol{r}}\boldsymbol{A}(\boldsymbol{0})_{(i,j)\in P}\hat{\boldsymbol{s}})\boldsymbol{i} - \left(\hat{\boldsymbol{r}}^{-1}\boldsymbol{A}(\boldsymbol{0})_{(i,j)\in N}\hat{\boldsymbol{s}}^{-1}\right)\boldsymbol{i} - \boldsymbol{v}(\boldsymbol{1})\right| < \varepsilon|\boldsymbol{v}(\boldsymbol{1})|$for arbitrarily small ε.

This generalized RAS variant was developed as an alternative to an ad hoc RAS approach in which negative entries are subtracted from the initial matrix and the balancing constraint in a first step. After balancing the resulting semi-positive matrix negative elements are added back to obtain the final matrix. As negative elements are ignored, they do not contribute to the minimization of information loss, yielding a sub-optimal solution. GRAS, however, incorporates negatives explicitly, so that information loss is lowered compared to the ad hoc approach [Junius & Oosterhaven, 2003].

The second major problem consists in the limitation to constraints on row and column sums only, which in our case prevents the introduction of additional information on single elements or on aggregates of elements of the MRSUT. The simplest approach for the introduction of additional information consists in excluding these elements from the balancing procedure. Afterwards these elements are added back to the solution of the conventional RAS balancing procedure [Miller & Blair, 2009]. As in many cases information about aggregates is known, several refinements of the RAS approach were developed e.g. handling constraints from national IOTs on multi-regional input-output tables. Oosterhaven et al. [1986 introduce an additional constraint to the minimal information loss problem that ensures consistency of intra and interregional transactions with national ones. Based on test for tables of three provinces they conclude that the use of national tables increases reliability of regional tables, especially of multi-regional tables compared to single regional ones. In the case that partial information is available on a higher level of aggregation than the original matrix to be balanced, a three-stage variant of RAS, the so called TRAS, may be used if aggregation rules are known [Gilchrist & St. Louis, 2004, Gilchrist & St Louis, 1999] ([Gilchrist & St Louis, 1999], [Gilchrist & St. Louis, 2004]). The first two steps consist of the conventional RAS in which adjustments towards known row and column sums are made, whereas in the third step adjustment factors are calculated by means of element wise division of the known matrix by the aggregated target matrix. These authors furthermore show that compared to the conventional RAS the introduction of additional information yields superior results [Gilchrist & St Louis, 1999]. Cole [1992], finally, proposes a general method for handling constraints of arbitrary size or shape on subsets of the matrix to be balanced. All constraints are thereby treated as subtotals of blocks of entries in a generalized formulation of constraints and are applied successive manner in each round of adjustment. Grounded on the economic interpretation of changes in input-output coefficients Snower [1990 proposed an updating method that is based on two sets of information about final demand and output as well as prices and value-added for the desired year of the updated table. Final demand is, then, used to estimate output in terms of the static quantity model using input-coefficients of the “old” matrix. Differences between estimated and actual output are said to be the result of systematic changes in input coefficients and are used to construct an adjustment vector for pre-multiplication. Information about value-added, however, is used to estimate a vector of relative prices, whose deviation from known prices is used to construct the second adjustment vector.

Third, the afore mentioned methods do not account for uncertainties neither of the initial estimate nor of external data for constraints. Hence, these methods do not distinct between elements of the initial estimate or constraints that are known with high certainty and highly uncertain elements. Recent attempts to address uncertainty of constraints on row and column total have inter alia been made by Daalgard and Gysting [2004] by introducing confidence factors that cause elements with high certainty to be adjusted more slowly than those that are highly uncertain during the adjustment process.

However, none of the methods are able to deal with inconsistencies in external data. Such inconsistencies can either have a direct or an indirect nature. Direct inconsistencies for example occur when data from different sources are imposed as constraints on one and the same element, but with different values. On the other hand, indirect inconsistencies arise if constraints on different elements lead to conflicts due to the interindustry relations present in input-output and supply-use tables. As such situations will any of the RAS approaches mentioned above prevent from converging inconsistencies have to be eliminated manually [Lenzen et al., 2009]. The recent RAS variant KRAS (where K stands for Konfliktfrei) was developed to solve the problem of conflicting data and is also capable to handle constraints of arbitrary size and shape, to handle negative negatives and to take uncertainty of constraints and the initial table into account. The main idea of the KRAS algorithm consists of adjusting values of constraints according to their uncertainty if the conventional GRAS terminates oscillating due to conflicts in constraint values. The formulation of the information loss problem is based on a vectorized representation $\boldsymbol{a}$ of the matrix $\boldsymbol{A}$ to be balanced or updated, such that constraints can be formulated as $\boldsymbol{Ga} = \boldsymbol{c}$, where $\boldsymbol{c}$ denotes RHS values of the constraints and $\boldsymbol{G}$is a matrix of constraint coefficientsg_{ij}. If it is for example the sum two elements is known, say $a_1 + a_2 = c_3$, then row 3 of $\boldsymbol{G}$consists of ones in column 1 and 2 and zeros elsewhere. If however c_3 is difference between both elements $(a_1 - a_2)$, then $g_{23} = -1$. Note, that an important feature of KRAS is that g_{ij} is not restricted to unity, as in some cases there is information on ratios rather than absolute values, e.g. the share of product output that is sold to households.

The corresponding Lagrangean appears as a modified version of that used in GRAS for allowing a_i to be positive as well as negative:

$$L = \sum_{i;a_i(0)\geq 0} a_i(1) \ln \frac{a_i(1)}{e\, a_i(0)} - \sum_{i;a_i(0)<0} a_i(1) \ln \frac{a_i(1)}{e\, a_i(0)} + \sum_i \lambda_i \left[\sum_i g_{ij} a_i(1) - c_i\right] \quad (64),$$

Where $a_i(0)$ and $a_i(1)$ are elements of the vectorized initial and final table, respectively. Working out the first order condition and solving for $a_k(1)$ yields following solution, where k is the index for constraints:

$$a_k(1) = a_k(0)e^{-\sum_i \lambda_i g_{ik}} = a_i(0) \prod_i e^{-\lambda_i g_{ik}} \qquad \text{for } a_k(0) \geq 0 \quad (65a)$$

$$a_k(1) = a_k(0)e^{\sum_i \lambda_i g_{ik}} = a_i(0) \prod_i e^{\lambda_i g_{ik}} \qquad \text{for } a_k(0) < 0. \quad (65b)$$

Inserting this solution into the constraint equation $\boldsymbol{Ga} = \boldsymbol{c}$delivers:

$$\sum_{j;a_j(0)\geq 0} g_{ij} a_j(0) \prod_k e^{-\lambda_k g_{kj}} + \sum_{j;a_j(0)<0} g_{ij} a_j(0) \prod_k e^{\lambda_k g_{kj}} - c_i = 0 \quad (66).$$

This equation is then solved iteratively starting with $\boldsymbol{a(0)} = \boldsymbol{a^{(0)}}$ arbitrarily chosen $\lambda_k{}^{(0)}$. For the purpose of solving inconsistencies in constraints, KRAS uses a modified termination condition, as in the case of conflicts the solution of conventional RAS as well as GRAS alternates between conflicting constraints. In the case of KRAS right hand side values of constraints are adjusted to solve conflicts if the distance between the constraint $\boldsymbol{c}$ and its realization $\boldsymbol{Ga}$ does not improve in two subsequent iterations n and (n-1):

$$|\boldsymbol{Ga} - \boldsymbol{c}|^{(\boldsymbol{n})} - |\boldsymbol{Ga} - \boldsymbol{c}|^{(\boldsymbol{n-1})} < \delta \quad (67)$$

If this inequality is fulfilled $\boldsymbol{c}$needsto be modified too, so that the iterative procedure is carried out in four steps, whereby $c_i(0) = c_i$:

1. With $0 \leq \alpha \leq 1$ and the standard deviation σ_i of c_i, compute $c_i{}^{(n)}$ according to
$$c_i{}^{(n)} = c_i{}^{(n-1)} - Sgn\left(c_i{}^{(n-1)} - \sum_j g_{ij} a_j{}^{(n-1)}\right) \times \min\left(\left|c_i{}^{(n-1)} - \sum_j g_{ij} a_j{}^{(n-1)}\right|, \alpha\sigma_i\right) \quad (68a)$$
2. For given $\lambda_i{}^{(n-1)}$and $c_i{}^{(n)}$solve
$$\sum_{j;a_j{}^{(n-1)}\geq 0} g_{ij} a_j{}^{(n-1)}\, e^{-(\lambda_i{}^{(n)} - \lambda_i{}^{(n-1)})g_{ij}} + \sum_{j;a_j{}^{(n-1)}\geq 0} g_{ij} a_j{}^{(n-1)}\, e^{(\lambda_i{}^{(n)} - \lambda_i{}^{(n-1)})g_{ij}} - c_i{}^{(n)} = 0 \qquad \text{for } \lambda_i{}^{(n)} \quad (68b)$$
3. With $\lambda_i{}^{(n)}$ given from step 1 compute.

$a_j{}^{(n)} = a_j{}^{(n-1)} e^{-(\lambda_i{}^{(n)} - \lambda_i{}^{(n-1)}) g_{ij}}$ for $a_j{}^{(n-1)} \geq 0$ or (68c)

$a_j{}^{(n)} = a_j{}^{(n-1)} e^{(\lambda_i{}^{(n)} - \lambda_i{}^{(n-1)}) g_{ij}}$ for $a_j{}^{(n-1)} < 0$ (68d)

4. With$c_i{}^{(n)}$,$\lambda_i{}^{(n)}$ and $a_j{}^{(n)}$ given return to step 1 and compute $\lambda_i{}^{(n+1)}$.

It becomes clear that the KRAS algorithm will violate constraints in the case of inconsistencies, but this violation is made based on uncertainty associated with that constraint. If $\left|c_i{}^{(n-1)} - \sum_j g_{ij} a_j{}^{(n-1)}\right| < \alpha\sigma_i$ equation 51 becomes $c_i{}^{(n)} = \sum_j g_{ij} a_j{}^{(n-1)} = c_i{}^{(n-1)}$ indicating that no adjustment is necessary. If on the other hand $\left|c_i{}^{(n-1)} - \sum_j g_{ij} a_j{}^{(n-1)}\right| < \alpha\sigma_i$ the constraint is adjusted according to standard deviation, so that KRAS finds a compromise solution between constraints in conflict [Lenzen et al., 2009]. In the case of conflicts between two constraints - e.g. c_1 and c_2, where c_1 is of higher confidence $\sigma_1 < \sigma_2$ - c_2 is modified more strongly than c_1, which results in a compromise solution closer to c_1.

The important refinement of the KRAS algorithm compared to other RAS variants consists in solving potential conflicts in external data in an automated manner. It, thus, removes the necessity of manual and sometimes arbitrary adjustments of data, but treads data uncertainty as a fundamental characteristic of real world data, that needs to be handled in a transparent manner.

IV.2 Processing of Superior Data

Data from national and regional accounts have already been used for the construction of the non-survey initial estimate. The introduction of these data items as additional constraints is necessary to ensure consistency of the final table with national and regional accounts. As the introduction of additional data during the balancing and adjustment process will have effects on regional and national aggregates (soft) boundaries for the adjustment will be set by regional and national accounts. As national and regional accounts use the same classification as the supply-use tables no reclassification or other adjustments were necessary. Regional foreign trade data for physical goods have to be reclassified from EWG to CPA classification, before they can use as constraints.

The remaining superior data namely transport flows and information from industrial cost structure survey, however, require several steps of estimation and processing. The methods and procedures applied to them will be explained in the next subsection more detailed.

IV.2.1 Freight Transport Data

Transportation data are available in in the form of complete OD-Matrices for 58 types of products for shipments in tons via rail and inland navigation. These data are complete, so that the data processing only consists of re-classification and the transformation from tons into monetary values as described below. However, for shipments via road OD-matrices are only available for 10 types of products, whereby there is only one complete OD-matrix for manufactured products. OD-matrices for other products often contain information on major flows, as well as zero flows and a "/" denoting positive flows, for which the standard error was to large. As total inflows and outflows of the German states are available for 20 different types of products in addition, the task consists in estimating individual flows between pairs of regions consistent with given information inflows and outflows as well as partial information on (sums of) individual flows. The estimation of individual flows is therefore in the context of spatial interaction models and it was decided to estimate OD-matrices for road transportation with an entropy maximizing model [Wilson, 1970].

IV.2.1.1 The Entropy Maximizing Approach

For this approach it was not possible to use inflow and outflow data for all 20 types of products for two reasons: First, for some products inflow and outflow data are incomplete at a deeper level of disaggregation. This was the case for non-solid fuels (crude oil and refined petroleum) as well as for products of agriculture and forestry, for which only the total of both is available. Second, some product groups exhibit a greater level of detail than the product classification of the SUTs. Chemical products for example comprise 4 different product subgroups. The only products group for which complete information on subgroups are available are manufactured products, comprising machinery,

Due to the nature of the road transportation data the entropy maximizing approach is used for the product-groups agriculture and forestry, food, non-solid fuels, basic metal, other non-metallic mineral products and chemical products to estimate missing values of the OD-matrix. On the other hand, for manufactured products comprising the subgroups transport equipment, fabricated metal and machinery, glass, textiles and wearing apparel, and other manufactured products, complete inflow and outflow data as well as a complete OD-matrix are available, so that the entropy maximizing approach is used to estimate OD-matrices for each subgroup. The results of these two procedures are 11 OD-matrices for road transportation in total. The structure of such OD-matrices is presented in table 6.

The basic idea of entropy maximization in the context of spatial interaction was developed by Wilson [1967 and is based on the concepts of micro and macro states of a spatial interaction system. The macro state of such a system is represented by the OD-matrix in table 6, which counts the number of transactions (in our case freight flows) between each possible pair of regions. The micro level of a spatial interaction system however consists of a description of the movement of each individual ton of freight from one region to another, so that each macro state is a representation of individual movements on the micro level.

Given a number of total number tons $T = \sum_r T_j^{r\cdot} = \sum_r T_j^{\cdot r}$ transported in the system, each micro state consistent with $T$is equally likely. From that follows that the probability of a specific macro state is proportional to the number of possible micro states that can reproduce that specific macro state in terms of aggregation. If for example in table 6 there are only shipments from region 1 to region 2 and shipments between all other regions are zero, then there is only one micro state that can reproduce this macro state and this is when each ton is shipped from region 1 to region 2. If, however, only one ton is shipped in the other direction, Tmicrostates exist that can yield the macro state (OD-matrix), as each single ton can be the one that is shipped in the other from region 1 to region 2 [Sargento, 2009, Snickars &

Weibull, 1977]. In general the number of micro states consistent with a macro state $E(\boldsymbol{T})$is given as:

$$E(\boldsymbol{T}) = \frac{T!}{\prod_r \prod_s t^{rs}} \quad (68).$$

The basic idea of the entropy maximizing approach is now to find the macro state (od-matrix) that is compatible with the maximum number of micro states and therefore the most probable distribution of interregional shipments [Wilson, 1970].

Table 6 Structure of OD transportation data

	Region 1	Region 2		Region r	Total
Region 1	t_j^{11}	t_j^{12}	$\cdots$	t_j^{1r}	$T_j^{1\cdot}$
Region 2	t_j^{21}	t_j^{22}	$\cdots$	t_j^{2r}	$T_j^{2\cdot}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
Region r	t_j^{r1}	t_j^{r2}	$\cdots$	t_j^{rr}	$T_j^{r\cdot}$
Total	$T_j^{\cdot 1}$	$T_j^{\cdot 2}$	$\cdots$	$T_j^{\cdot r}$	$\sum_r T_j^{r\cdot} = \sum_r T_j^{\cdot r}$

Source: own calculations IEK-STE 2013

Due to this the entropy maximizing approach belongs to the class of probabilistic spatial interaction models. For using (48) as an objective function a logarithmic transformation is often applied on $E(\boldsymbol{T})$, which yields:

$$\ln E(\boldsymbol{T}) = \ln T! - \sum_r \sum_s \ln t^{rs}! \quad (70).$$

The constant $\ln T!$ can be eliminated for the maximization and Stirling's approximation[10] can be used for the factorials, so that the objective function for the entropy maximizing model for product j appears as:

[10] $\ln x! = x(\ln x - 1)$, see Wilson [1967.

$$\max_{t_j^{rs}} \ln E(\boldsymbol{T}) = -\sum_r \sum_s \left[t_j^{rs} \ln\left(t_j^{rs}\right) - t_j^{rs} \right] \qquad (71).$$

The example above illustrates, that the unconstraint maximization of (71) yields an even distribution of shipments, so that additional information should be used to constrain the solution. In this case the distribution of shipments must be such that given row and column sums are met, i.e. $\sum_r t_j^{rs} = T_j^{\cdot s}$ and $\sum_s t_j^{rs} = T_j^{r\cdot}$. In this case the entropy maximizing model is called double constraint [Wilson, 1971].

For the application of entropy maximizing models it is furthermore useful to make use of some sort of transportation cost for shipments between the pairs of regions by product as well as total costs of shipments by product. As transport margins are reported neither in the official German SUTs nor in the WIOD database, it is assumed that costs are proportional to the average transportation distance between regions r and s $dist^{rs}$. In addition to tons inflows and outflows are also available in terms of ton kilometers $TK_j^{\cdot s}$ and $TK_j^{r\cdot}$ by product, which is the product of tons of product j entering or leaving region r and the transport distance. From this data two transport cost constraints for each region and product can be constructed:

$$\sum_r t_j^{rs} * dist^{rs} = TK_j^{\cdot s} \qquad (72a)$$

$$\sum_s t_j^{rs} * dist^{rs} = TK_j^{r\cdot} \qquad (72b)$$

Following Nitsch [2000] distances are estimated as weighted distances from great circle distances between the capitals of more than 400 counties and from population data as weighting factors. This approach was originally used by Helliwell and Verdier [2001] and further refined by Nitsch for the estimation of intranational trading distances for empirical studies on the border effect. Intrastate trading distances are estimated as:

$$dist^{rr} = \frac{\sum_n \left(\sqrt{\pi}^{-1} \sqrt{area_n}\ pop_n^2\right) + \sum_{n \neq m} (pop_n\ \ pop_m\ \ dist^{nm})}{\sum_{nm} pop_n\ \ pop_m} \qquad (73),$$

where n and m are counties belonging to state r , pop_n denotes the population in county n and $area_n$ the area of county n. The first term of the numerator approximates the average intracounty trading distance as a function of land area multiplied by $\sqrt{\pi}^{-1}$, which comes from the approximation of county's shape as a circle. The second term is the intercounty distance weighted by their population.

The average trading distance of pair of German states is estimated as follows, whereby n and m denote counties belonging the states r and s respectively:

$$dist^{rs} = \frac{\sum_{nm} (pop_n\ \ pop_m\ \ dist^{nm})}{\sum_{nm} pop_n\ \ pop_m} \qquad (74).$$

Given these data for shipments and tons, ton*kilometers and distances the entropy maximization problem for the estimation of missing values of the OD-matrices for product j can be stated as:

$$\max_{t_j^{rs}} E = -\sum_r \sum_s \left(t_j^{rs} \ln\left(t_j^{rs}\right) - t_j^{rs}\right) \qquad (75a)$$

s.t.

$\sum_r t_j^{rs} = T_j^{\cdot s}$ (75b)

$\sum_s t_j^{rs} = T_j^{r\cdot}$ (75c)

$\sum_r t_j^{rs} * dist^{rs} = TK_j^{\cdot s}$ (75d)

$\sum_s t_j^{rs} * dist^{rs} = TK_j^{r\cdot}$ (75e)

The constraints (75b) and (75c) ensure that shipments from region r to region s are consistent with $T_j^{\cdot s}$total inflows of region s and with total outflows of region r $T_j^{r\cdot}$.

In the case of manufactured products the elements of existing OD-matrix t_j^{rs} can be interpreted as capacity of arcs of a network connecting region r with region s [Ahuja et al., 1993]. For each product subgroup a positive parameter α_p^{rs} is defined, which denotes the share of capacity t_j^{rs} consumed by product subgroup p. The entropy maximizing problem may then be stated as:

$\max_{\alpha_p^{rs}} E = -\sum_p \sum_r \sum_s \left[\alpha_p^{rs} t_j^{rs} \ln\left(\alpha_p^{rs} t_j^{rs}\right) - \alpha_p^{rs} t_j^{rs}\right]$ (76a)

s.t.

$\sum_r \alpha_p^{rs}\, t_j^{rs} = T_p^{\cdot s}$ (76b)

$\sum_s \alpha_p^{rs}\, t_j^{rs} = T_p^{r\cdot}$ (76c)

$\sum_r \alpha_p^{rs}\, t_j^{rs} * dist^{rs} = TK_p^{\cdot s}$ (76d)

$\sum_s \alpha_p^{rs} t_j^{rs} * dist^{rs} = TK_p^{r\cdot}$ (76e)

$\sum_p \alpha_p^{rs}\, t_p^{rs} = t_j^{rs}$ (76f)

The additional bundle constraint (76f) ensures consistency of subgroup-flows with prior information of the OD-matrix for manufactured products.

IV.2.1.2 Conversion into Monetary Values

The final step of the estimation of constraints for interregional trade consists of the transformation of shipments from tons to monetary values. According to Llano et al. [2010] value to ton ratios can be calculated from regional export data, which are in Germany available at a 4-digit EWG classification, and applied to all outflows of the respective state. It is, thus, assumed that there are no differences in value to ton ratios with respect to modes of transportation and regions of destination. More precisely, as each product-group consists of many different types and qualities of products, it is implicitly assumed that the composition of each product group in interregional trade is the same as in foreign trade.

Despite of requiring strong assumptions about the internal composition of product groups this approach, furthermore, ignores some important relationships between the average prices per ton of products, the physical weight and the distance of transport-flows. In foreign trade statistics it can be observed for a product-group that high ton flows are associated with low average prices per ton, whereas flows over long distances are associated with high prices per ton.

With respect to distance, Baldwin and Harrigan [2011 observe this effect in f.o.b. unit-values of US exports showing this observation to be inconsistent with standard models of international trade such as those purely based on comparative advantage [Eaton & Kortum, 2002], a multi-country version of the Helpman and Krugman [1985 model with trade cost and a multi-country Melitz [2003 model. The same relationship between prices per ton and distance is reported by Helble and Okubo [2006 and Kneller and Yu [2008 for data on Chinese exports. Baldwin and Harrigan introduce product quality into the heterogeneous firms Melitz [2003 model and show that expansive high quality products are more competitive on distant markets and more likely to overcome distance related trade barriers. Johnson [2012 interprets this as Alchian-Allen Effect, since in the case of two substitute goods (high and low quality) fixed per ton transport costs decrease relative prices of high quality products making them more competitive.

The relationship between the amounts of tons ship and unit-values also reflects the composition of individual products within a product group in relation to origin and destination of an flow observed. As high ton flows take place within a region or between two neighboring regions, they most likely consist of a higher share of intermediate products (with low unit-values compared to finished goods of that product group) due to the existence of regional clusters and vertical integration of industries. Furthermore, transport economics suggest that products with low unit-values such as building materials or cereals are transported over shorter distances than those with high unit values like machines or cars. The share of transport costs in purchaser prices is higher for products with low unit-values, which makes them less competitive compared to locally produced commodities.

The relationship between the monetary values of shipments, its physical weight and distance can be expressed as follows:

$$\hat{t}_j^{rq} = \propto {t_j^{rq}}^{\beta_1} {dist^{rq}}^{\beta_2} \qquad (78),$$

where $\hat{t}_j^{rq}$ denotes the monetary value of exports of product j from state r to country q, t_j^{rq} denotes the physical weight of these exports, $dist^{rq}$ denotes the distance between a region/country pair and $\propto$, β_1 and β_2 are coefficients. Dividing (78) by t_j^{rq} yields the relationship between unit-values p_j^{rq}, ton-flows and distance:

$$\hat{t}_j^{rq}/t_j^{rq} = p_j^{rq} = \propto {t_j^{rq}}^{\beta_1 - 1} {dist^{rq}}^{\beta_2} \qquad (78)$$

For $\beta_1 < 1$ larger ton flows yield lower average prices per ton than smaller flows, whereas shipments over long distances are associated with higher unit-values compared to shipments over short distances.

In sake for exploring and quantifying these relationships data on exports of the German federal states to 41 European countries for 30 product-groups (GP03 Classification) are used. These 30 products are then further aggregated to match with product classification of transport statistics as well as with the supply-use tables. Hence, for monetary values as well as for physical weight the sample consists of 656 individual flows per product. Distances between German states and European countries were calculated via Google maps, whereby the program calculates the distances between geographical centers. The relationship between monetary values, physical weight and distance as expressed in equation (78) is then estimated in its logarithmic form:

$$\ln \hat{t}_j^{rq} = \alpha + \beta_1 \ln t_j^{rq} + \beta_2 \ln dist^{rq} + \varepsilon_j \quad (79)$$

and ε_j is an i.i.d error term. (51) was estimated in terms of stepwise regression.

The results of the most successful specifications for each type of product are shown in table 7. For each type of products β_1 and β_2 are significant on the 1% level, except the effect of distance on the monetary value of ore exports. The intercept α however is only significant for five product groups (machines and transport equipment).

The estimation result for β_1 shows that $0 < \beta_1 < 1$, which is consistent with expectation that large ton flows are associated with lower unit-values. Furthermore, β_1 is smaller for commodities with low unit values compared to those products with high prices per ton, which indicates that these products are transported over rather small distances. In the case of basic and fabricated metals for example the effect of tons transported on average prices of flows is stronger for basic metals than for fabricated ones. The same can be observed comparing agricultural with products such as machines or motor vehicles.

With respect to the effect of distance on the monetary value of exports, all coefficients (with an exception for ore) are $0 < \beta_2 < 1$, indicating that unit-values of shipments over long distances are higher than those for small distances for the same product group. This result is in line with results from international trade literature and the interpretation that unit-values of exports are subject to an Alchian-Allen Effect, as suggested by Johnson [2012. Only in the cases of crude oil/natural gas and coal distance has no effect on unit-values.

As distance between two trading partners can also be expected to have explanatory power for the amount of tons shipped from one region to the other, variance inflation factors (VIF) were estimated, in order to analyze the degree of potential multicollinearity. The variance inflation factors are calculated as:

$$VIF_j = \frac{1}{(1-\hat{R}_j^2)} \quad (80)$$

In formula (80) $\hat{R}_j^2$ denotes the coefficient of determination of an auxiliary regression of distance on shipments in tons: $\ln t_j^{rq} = \alpha + \beta_1 \ln dist^{rq} + \varepsilon_j$. VIFs as reported in table 7 are far away from critical values, indicating serious problems with multicollinearity (see O'brien [2007). Surprisingly, distance is of only little explanatory power for the amount of tons shipped between region/country pairs for all products, suggesting that a large share of deviations in monetary trade flows is caused by the effect of distance on unit-values rather than on physical weight.

Table 7 Results from stepwise Regression

Product	Coefficients			p-value			Adj. R^2	VIF
	α	β_1	β_2	α	β_1	β_2		
Agriculture/Forestry	-	0.7597	0.2475	-	0.0000	0.0000	0.8513	1.1077
Coal	-	0.7632	-	-	0.0000	-	0.9583	1.1826
Crude Oil	-	0.8948	-	-	0.0000	-	0.9941	1.0308
Ore	2.1383	0.8343	-0.2963	0.0877	0.0000	0.0991	0.8259	1.0009
Other mining	-	0.7203	0.0358	-	0.0000	0.0026	0.8991	1.1472
Food/Tobacco	-	0.8593	0.6326	-	0.0000	0.0000	0.2982	1.0855
Textile Products	0.9560	0.8947	0.2479	0.0927	0.0000	0.0012	0.8858	1.0947
Refined Petroleum	-	0.8413	0.1837	-	0.0000	0.0000	0.9731	1.1167
Chemical Products	-	0.8838	0.2856	-	0.0000	0.0000	0.9080	1.0916
Glass/Building Mt.	-	0.7600	0.2635	-	0.0000	0.0000	0.8912	1.1753
Basic Metals	-	0.8477	0.2912	-	0.0000	0.0000	0.9265	1.1111
Fabricated Metals	1.2004	0.9031	0.1629	0.0058	0.0000	0.0045	0.9397	1.1298
Machinery	2.6309	0.9133	0.1050	0.0000	0.0000	0.0426	0.9388	1.0777
Motor Vehicles	1.5019	0.9975	0.0999	0.0006	0.0000	0.0809	0.9375	1.0794
Other Manufact.	-	0.8942	0.2213	-	0.0000	0.0000	0.9373	1.1342

Source: own calculations IEK-STE 2013

Finally, the original equation (78) suggests a Cobb-Douglas type relationship, but for all products except agriculture/forestry the Null-Hypothesis $\beta_1 + \beta_2 - 1 = 0$ must be rejected in Wald-Tests.

For the estimation of interregional trade the interregional transport flows are transformed from tons into monetary units by applying (78) on interregional ton flows and distances. As there are large deviations between total outflows and inflows from regional output and regional consumption, it was decided to use the interregional trade-flows in monetary units as column ratio constraints.

IV.2.2 Cost Structure Survey

The construction of constraints on regional cost and output structures by industry is based on survey information collected from establishments and enterprises of the mining and manufacturing sectors. Survey information consists of two data sets; this is cost structure surveys on enterprises and annual reports of enterprises.

The variables of the cost structure survey for mining and manufacturing is presented in table 6 and the variables of the annual report in table 7 in the appendix. The major problem in the processing of these data is that many enterprises consist of several establishments that are located in different regions, i.e. multi-regional enterprises. Simple aggregation of data for industries and regions would, therefore, deliver biased results. Thus, the regional allocation of cost is achieved with help of the annual report, as the identification of establishments belonging to an enterprise is possible due to the enterprise IDs. For enterprises all types costs are allocated on establishments according to establishment's share in turnover of the respective enterprise. Afterwards, costs allocated to establishments are aggregated with respect to the regional location and the belonging to a 2-digit NACE Rev. 1.1 industry of the enterprise. For the regional supply tables it is assumed that each establishment produces only products that belong to the same 2-digit CPA category. Establishment data are then aggregated with respect to location, belonging of the establishments to an industry and belonging enterprises to an industry. Due to this assumption it is possible to interpret the turnover of an establishment as turnover from a specific product j.

There is also information on the structure of turnover from different business activities of enterprises, which can be used for constraints on regional supply tables. The same procedure as for costs is, therefore, also applied to different types of turnover.

For example an enterprise in Saxony belonging to the food industry consists of two establishments, whereby the first establishment produces food is located in Saxony and achieves 60% of total turnover. If the second establishment is located in Bavaria and produces chemical products, then 60% of costs are allocated to the food industry in Saxony's use table and 40% is allocated to the food industry in Bavaria's use table. On the other hand 60% of the enterprise's output is allocated to food production of the food industry in Saxony, whereas the remaining 40% is allocated to output of chemical products of the food industry in Bavaria.

Unfortunately, it is not possible to use the whole information of the cost-structure survey, as some categories of intermediate consumption and value-added are more detailed in the survey than in the use tables. For this reason some of the cost categories had had to be aggregated, so that finally categories of value-added (wages, depreciation, net operating surplus and net taxes on production) and five types of intermediate consumption are reached (material inputs, energy, renting of buildings and equipment, insurance services and other services). With respect to turnover it is possible to distinguish turnover from physical goods, trade services and other services. Note, that for the calculation of output, commodities purchased for re-sale ne to be subtracted from turnover generated by trade activities, as output of trade services only consists of the trade margin [Reich et al., 1995].

Before cost-structure information can be introduced as constraints, data need to be adjusted in two further steps, because the survey only covers enterprises with more than 20 employees and because purchases are reported at purchaser prices. The first task is addressed by calculating shares of costs and turnover in value-added for each industry and region. These ratios are, then, applied to industry specific value-added data given by regional accounts assuming small business enterprises to have average cost and turnover structure as those enterprises, who participated in the survey.

For the conversion of purchaser into basic prices is carried out analogue to the conversion of final consumption of households. It is assumed that industry specific net tax rates and trade margins as reported in the WIOD valuation matrices also apply to respective industries on the regional level. The sum of net taxes deduced by industry are then used as constraints on net product taxes, whereas the sum trade margins deduced is used as a constraint on intermediate consumption of trade services. Finally, the resulting information about primary and intermediate consumption (at basic prices) as well as turnover structures is used for ratio constraints.

IV.2.3 Estimation of Standard Deviations

A crucial requirement for using KRAS for balancing the German MRSUT is the availability of information about data uncertainty. Apart from being required for finding compromise solutions in the case of conflicting data, uncertainty information also play an important role for the assessment of reliability of results from input-output analysis based in GerMRIO. For the estimation of data uncertainties the methodology of Wiedmann et al. [2008] was adopted, which relies on the assumption of log-normally distributed data values. Then, the "order-of-magnitude errors" of a data value x can be approximated as:

$$\Delta\log_{10}(x) \approx \log_{10}(1 + r_x) \quad (81),$$

where $r_x = \Delta x / x$ is the relative standard error. The estimation of data uncertainty is based on the observation that large data values are generally more certain than smaller ones, as they usually consist of many small observation points such that errors are cancelled out during accumulation. For this reason following equation is estimated using ordinary least squares:

$$\log_{10}(1 + r_x) = \alpha + \beta * \ln(|x|) + \varepsilon \quad (81),$$

where α and β are regression coefficients to be estimated and ε is an error term. As no published information from which relative standard error can be constructed for the regression, differences between the values of different data sources are used as a proxy. For uncertainty inherent in the national supply-use tables the differences of the elements between the WIOD SUT and the official one are used as a proxy. Whereas for regional foreign trade and the cost structure survey data points were aggregated to national values and compared with information from the official supply-use table. Since for interregional trade none of such benchmark data exist minimum, maximum and average standard errors are set manually to values exceeding those for other constraint. This approach at least assures that interregional trade information is indicated as the most uncertain data item.

V Implementation and Results

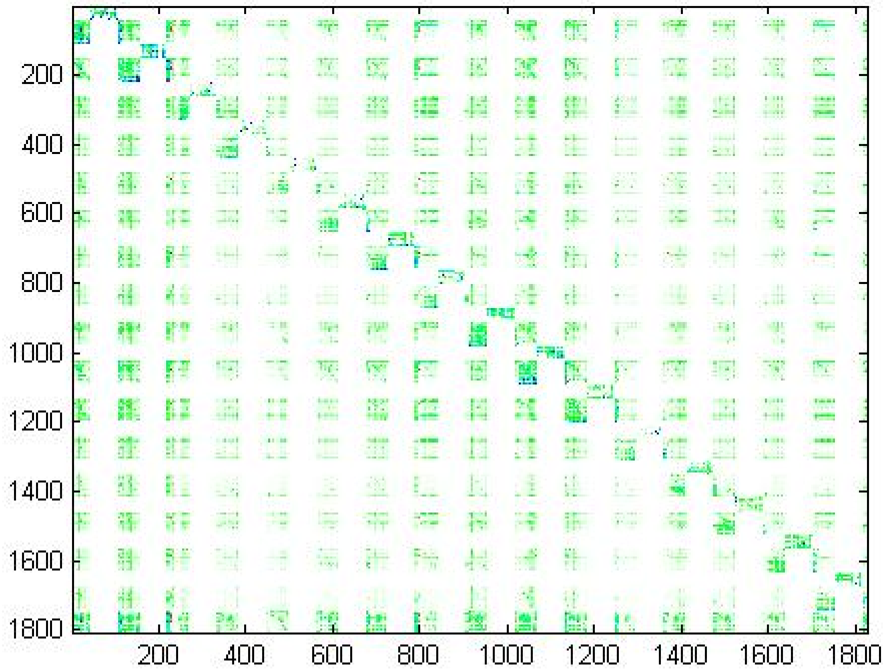
In order to generalize and facilitate the import of the initial estimate and additional data, as well as the formulation of constraints, an 8-tiered hierarchy, the so-called tree-structure of the table, is used to address single or sets of elements. The first two tiers describe the accounting year and the valuation of the table, which is in our case each only consist of one element (2007 and basic prices). The following three tears identify the origin of transactions reported in the table by specifying region, entity and sector of origin, whereas the remaining three tiers specifies region, entity and sector of destination. The German MRSUT incorporates 17 regions (16 federal states + row) and 4 entities (industries, commodities, value-added and final demand), each of them consisting of several sectors. In total there are 41 industries, 63 Commodities, 5 components of value-added and 10 final demand sectors.

For the formulation of the actual optimization problem initial and final MRIO as well as raw data are vectorised automatically. The $N \times 1$ vectorised final MRIO$\boldsymbol{P}$must, then, fulfil the constraints of raw data expressed as a system of linear equations $\boldsymbol{GP} = \boldsymbol{c}$ and of a set of boundaries $\boldsymbol{l} \leq \boldsymbol{P} \leq \boldsymbol{u}$ as much as possible, whereby $\boldsymbol{c}$is the $M \times 1$column vector of raw data and $\boldsymbol{G}$is a $M \times N$ matrix of coefficients. $\boldsymbol{l}$and $\boldsymbol{u}$are lower and upper boundaries. The boundaries $\boldsymbol{l}$and$\boldsymbol{u}$allows to pose restriction on the sign of elements. The vast majority is required to be non-negative ($l = 0$ and $u = +\infty$), such as output, intermediate- or final consumption. Only few elements are allowed to become negative, these are changes in inventories, net taxes on products and production and net operating surplus. In fact the industrial cost-structure survey revealed negative net profits of some regional manufacturing industries. The coefficient matrix $\boldsymbol{G}$is used to connect elements of the MRIO to raw data points. If for example the sum of a set of MRIO elements (e.g. the sum of total steel used by machinery industry of all regions) is required to meet specific value (e.g. value of steel used by the machinery industry of the national use table), the respective row of $\boldsymbol{G}$will be populated with zeros and ones. In cases where raw data is more/less aggregated as in the MRIO or have different classifications, operations of aggregation, disaggregation and reclassification are conducted automatically by making use of concordance matrices [Geschke et al., 2011].

For the operation of AISHA the data processing language A-Lang is used to read and locate sections of the initial estimate and to pose constraints in terms of raw data on them. Each AISHA-run requires at least three types of command files: First, the tree-structure file contains the structure of the MRIO in terms of the 8-dim representation, i.e. it defines regions, entities and sectors. The second file contains A-Lang commands that read the initial estimate and identify the positions of individual sections in the MRIO accounting framework. Finally, the constraint-file consists of A-Lang commands that steer the insertion and processing of raw data and formulates their relation to the MRIO in terms of constraints [Geschke et al., 2011].

The result of the final table can be visualized in various ways. Figure 2 shows a so-called heat-map, which gives an overview over the structure and values of the MRIO. The scale ranges from dark blue for strongly positive to red for negative elements. Diagonal blocks represent regional supply-and intraregional use tables, whereas off-diagonal blocks represent interregional trade transactions according to the use-regionalized framework.

Figure 2 Heatmap of the German MRSUT

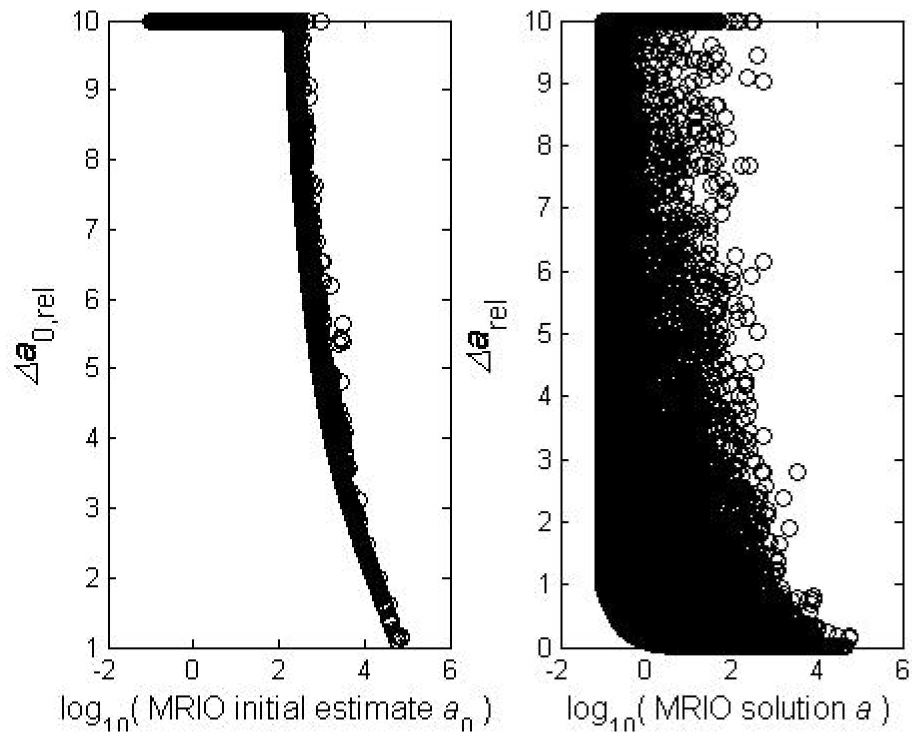


Source: own calculations IEK-STE 2013

It can be seen even on the regional level, where trade-borders are absent, intraregional transaction are still more important compared to interregional trade in most cases. An exception are especially the city states (Berlin, Hamburg and Bremen), which have a very low level local industrial production and have a very high dependency from surrounding regions.

Another important aspect is the visualization of data uncertainty, i.e. the adherences of MRIO values from raw data constraints. The comparison of relative adherences is shown in Figure 3. The y-axis displays the relative violation of constraints and the x-axis the logarithm of MRIO elements. It can be seen from the left diagram that the elements of the initial estimate with low values violate raw data constraints much more strongly than elements with high values. Overall, there is a strong relationship between relative adherences and the size of elements. The left diagram however shows the adherences from raw data of the final MRSUT and gives a good impression on how KRAS works: elements at the bottom are those, which are adequately supported by raw data with low uncertainty. It can also be seen that elements with high values have much lower uncertainty (bottom right), i.e. for elements with a magnitude of more than four relative violations of raw data are close to zero. Cases in which elements of relatively high magnitude (top) strongly violate raw data are seldom, which indicates that the vast majority of these elements are adequately supported by data. Overall, total adherences of the final table are significantly minimized compared to the initial values. This result encourages concluding that the final table is accurate in terms of holistic accuracy.

Figure 3 relative adherences of initial and final MRIO elements from constraints



Source: own calculations IEK-STE 2013

VI Conclusion

The focus of this report has been the construction procedure of German multi-regional supply-use table. It was demonstrated how the integration of initial estimate and excessive amounts of superior data can simplified and enhanced, due to standardisation and automation coming along with the AISHA software package. Thereby, a crucial aspect of the construction is the transparent handling and reporting of data uncertainty, which allows the critical evaluation of results gained from analysis. The MRSUT in its current form is well equipped for the analysis of effect of energy policy on regional structural change and spatial as well as social distribution of gains and burdens.

Based on the MRSUT multi-regional interindustry model will developed in three steps, whereby each sub-model is applied to a specific issue of the energy-policy: 1.) Static quantity model with endogenous consumption of households for the analysis of regional distribution and spill-over effects from the promotion of renewable energies. 2.) Static price model for studying effects of feed-in tariffs on regional economies and income distribution. 3.) Introduction of dynamics to the model in terms of endogenous investment and analysis of regional energy policies in comparison to national energy policy in the long-run. Apart from AISHA allows the extension of the German MRSUT database in several directions in a consistent and timesaving manner:

First, the German MRSUT can be connected to global or European input-output databases such as EORA or WIOD, in order to become a mixed interregional/international database. This extension would greatly enlarge the range of analysis, as it allows analysing effects of global developments (in certain countries) on regional energy supply in Germany or the shift of production to foreign countries due to national policies. A second possible extension is the further disaggregation of certain important sectors and the introduction of data about energy use in physical units and greenhouse gas emissions. In this way, macroeconomic consequences of efficiency improvements in certain sectors, e.g. steel production can be studied and the possibilities of connecting input-output to technology oriented models are increased. The combination of both directions offers additional potential applications such as calculating regional carbon footprints or hybrid LCA, which consists of connecting standard life cycle to input-output analysis.

In summary, the German MRSUT offers a comprehensive database and a flexible framework for various applications, especially for the evaluation of energy policy, so that this tool could improve insights of researches as well as the support of political decision-making.

VII References

AHUJA, R. K., MAGNANTI, T. L. & ORLIN, J. B. (1993) Network flows: theory, algorithms, and applications.

ALMON, C. (2000) Product-to-product tables via product-technology with no negative flows. *Economic Systems Research,* 12:1**,** 27-43.

BACHARACH, M. (1970) *Biproportional matrices and input-output change*, CUP Archive.

BALDWIN, R. & HARRIGAN, J. (2011) Zeros, Quality, and Space: Trade Theory and Trade Evidence. *American Economic Journal: Microeconomics,* 3:2**,** 60-88.

BONET, J. (2005) Regional Structural Changes in Colombia: An Input-Output Approach. *Colombia: Banco de la Repùblica de Colombia. Borradores de Economia*:341.

BONFIGLIO, A. & CHELLI, F. (2008) Assessing the behaviour of non-survey methods for constructing regional input–output tables through a Monte Carlo simulation. *Economic Systems Research,* 20:3**,** 243-258.

BOUWMEESTER, M. C. & OOSTERHAVEN, J. (2008) *Methodology for the construction of an international supply-use table*. International Input-Output Meeting on Managing the Environment.

CHENERY, H. B. (1953) *Regional Analysis*. In CHENERY, H. B. (Ed.) *The Structure and Growth of the Italien Economy.* Rome, J. S. Mutual Security Agency.

COLE, S. (1992) A note on a Lagrangian derivation of a general multi-propotional scaling algorithm. *Regional Science and Urban Economics,* 22:2**,** 291-297.

DALGAARD, E. & GYSTING, C. (2004) An algorithm for balancing commodity-flow systems. *Economic Systems Research,* 16:2**,** 169-190.

DANIELS, P. L., LENZEN, M. & KENWAY, S. J. (2011) THE INS AND OUTS OF WATER USE – A REVIEW OF MULTI-REGION INPUT–OUTPUT ANALYSIS AND WATER FOOTPRINTS FOR REGIONAL SUSTAINABILITY ANALYSIS AND POLICY. *Economic Systems Research,* 23:4**,** 353-370.

DE MESNARD, L. (2004) Understanding the shortcomings of commodity-based technology in input-output models: an economic-circuit approach. *Journal of Regional Science* 44:1**,** 125-141.

DIVISION", U. N. S. (2008) *International Merchandise Trade Statistics: Supplement to the Compilers Manual*, United Nations Publications.

EATON, J. & KORTUM, S. (2002) Technology, geography, and trade. *Econometrica,* 70:5**,** 1741-1779.

EDING, G., OOSTERHAVEN, J., DE VET, B. & NIJMEIJER, H. (1999) *Constructing Regional Supply and Use Tables: Dutch Experiences*. In HEWINGS, G., SONIS, M., MADDEN, M. & KIMURA, Y. (Eds.) *Understanding and interpreting economic structure.*237, Berlin Heidelberg New York, Springer.

EUROSTAT (1996) European System of Accounts, ESA. http://circa.europa.eu/irc/dsis/nfaccount/info/data/ESA95/en/titelen.htm.

EUROSTAT/EUROPEAN COMMISSION (2008) *Eurostat Manuel of Supply, Use and Input-Output Tables.* Office for Official Publications of the European

Communities, Luxembourg. http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-RA-07-013/EN/KS-RA-07-013-EN.PDF.

FLEGG, A. T. & TOHMO, T. (2011) Regional Input–Output Tables and the FLQ Formula: A Case Study of Finland. *Regional Studies,* 47:5**,** 703-721.

GALLEGO, B. & LENZEN, M. (2009) *Estimating Generalised Regional Input–Output Systems: A Case Study of Australia*. In RUTH, M. & DAVIDSDOTTIER, B. (Eds.) *The dynamics of regions and networks in industrial ecosystems.*55-82, Cheltenham, Edward Elgar Publishing.

GESCHKE, A., LENZEN, M., KANEMOTO, K. & MORAN, D. (2011) *AISHA: A tool to construct a series of contingency tables*. 19th IIOA Conference, Alexandria, US.

GILCHRIST, D. & ST. LOUIS, L. (2004) An Algorithm for the Consistent Inclusion of Partial Information in the Revision of Input-Output Tables. *Economic Systems Research,* 16:2**,** 149-156.

GILCHRIST, D. A. & ST LOUIS, L. V. (1999) Completing Input–Output Tables using Partial Information, with an Application to Canadian Data. *Economic Systems Research,* 11:2**,** 185-194.

GORDON, P., MOORE, J. E., PARK, J. Y. & RICHARDSON, H. W. (2007) The economic impacts of a terrorist attack on the US commercial aviation system. *Risk Analysis,* 27:3**,** 505-512.

HELBLE, M. & OKUBO, T. (2006) Heterogeneous quality and trade costs. *Unpublished, Graduate Institute of International Studies*.

HELLIWELL, J. F. & VERDIER, G. (2001) Measuring internal trade distances: a new method applied to estimate provincial border effects in Canada. *Canadian Journal of Economics***,** 1024-1041.

HELPMAN, E. & KRUGMAN, P. R. (1985) *Market structure and foreign trade: Increasing returns, imperfect competition and the international economy*, The MIT press.

HOFFMAN, R. B. K., J.N. (1974) *Design for Commodity-by-Industry Interregional Input-Output Models.* Cambridge, Ballinger.

HULU, E. & HEWINGS, G. J. (1993) THE DEVELOPMENT AND USE OF INTERREGIONAL INPUT-OUTPUT MODELS FOR INDONESIA UNDER CONDITIONS OF LIMITED INFORMATION*. *Review of Urban & Regional Development Studies,* 5:2**,** 135-153.

ISARD, W. (1951) Interregional and regional input-output analysis: a model of a space-economy. *The review of Economics and Statistics,* 33:4**,** 318-328.

ISARD, W., AZIS, I. J., DRENNAN, M. P., MILLER, R. E., SALTZMAN, S. & THORBECKE, E. (1998) *Methods of interregional and regional analysis*, Ashgate Aldershot.

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.

JACKSON, R. W. (1998) Regionalizing national commodity-by-industry accounts. *Economic Systems Research,* 10:3**,** 223-238.

JACKSON, R. W. & SCHWARM, W. R. (2011) Accounting foundations for interregional commodity-by-industry input-output models. *Letters in Spatial and Resource Sciences,* 4:3**,** 187-196.

JANSEN, P. K. & RAA, T. T. (1990) The choice of model in the construction of input-output coefficients matrices. *International Economic Review***,** 213-227.

JENSEN-BUTLER, C. & MADSEN, B. (1996) Modelling the regional economic effects of the Danish Great Belt link. *Papers in Regional Science,* 75:1**,** 1-21.

JENSEN, R. C. (1980) The concept of accuracy in regional input-output models. *International Regional Science Review,* 5:2**,** 139-154.

JOHNSON, R. C. (2012) Trade and prices with heterogeneous firms. *Journal of International Economics,* 86:1**,** 43-56.

JUNIUS, T. & OOSTERHAVEN, J. (2003) The solution of updating or regionalizing a matrix with both positive and negative entries. *Economic Systems Research,* 15:1**,** 87-96.

KNELLER, R. & YU, Z. (2008) Quality selection, Chinese exports and theories of heterogeneous firm trade. *University of Nottingham, GEP Research Paper,* 44.

KRONENBERG, T. (2009) Construction of Regional Input-Output Tables Using Nonsurvey Methods The Role of Cross-Hauling. *International Regional Science Review,* 32:1**,** 40-64.

KRONENBERG, T. & TÖBBEN, J. (2011) Regional input-output modelling in Germany: The case of North Rhine-Westphalia.

LAHR, M. L. (1993) A review of the literature supporting the hybrid approach to constructing regional input–output models. *Economic Systems Research,* 5:3**,** 277-293.

LAHR, M. L. (2001) Reconciling domestication techniques, the notion of re-exports and some comments on regional accounting. *Economic Systems Research,* 13:2**,** 165-179.

LEHMANN, H. (2004) *Die Modellierung der Konsumausgaben privater Haushalte: auf der Grundlage repräsentativer Einkommens-und Verbrauchsstatistiken*, Nomos Verlag-Ges.

LEIGH, R. (1970) The use of location quotients in urban economic base studies. *Land Economics,* 46:2**,** 202-205.

LENZEN, M., GALLEGO, B. & WOOD, R. (2009) Matrix balancing under conflicting information. *Economic Systems Research,* 21:1**,** 23-44.

LENZEN, M., MORAN, D., KANEMOTO, K., FORAN, B., LOBEFARO, L. & GESCHKE, A. (2012) International trade drives biodiversity threats in developing nations. *Nature,* 486:7401**,** 109-112. 10.1038/nature11145

LENZEN, M., MORAN, D., KANEMOTO, K. & GESCHKE, A. (2013) Building EORA: A Global Multi-Region Input–Output Database at High Country and Sector Resolution. *Economic Systems Research,* 25:1**,** 20-49.

LEONTIEF, W. S., A. (1963) *Multi-regional Input-Output Analysis.* New York, St. Martin's Press.

LIANG, Q.-M., FAN, Y. & WEI, Y.-M. (2007) Multi-regional input–output model for regional energy requirements and CO_2 emissions in China. *Energy Policy,* 35:3**,** 1685-1700.

LLANO, C., ESTEBAN, A., PÉREZ, J. & PULIDO, A. (2010) Opening the Interregional Trade ‘‘Black Box’’: The C-Intereg Database for the Spanish Economy (1995—2005). *International Regional Science Review,* 33:3**,** 302-337.

LOUHELA, T. (2006) *Estimation of interregional trade for Finnish regions in 1996 and 2002–freight flow and gravity approaches*. 46th Congress of the European Regional Science Association, Volos, Greece.

MADSEN, B. & JENSEN-BUTLER, C. (1999) Make and use approaches to regional and interregional accounts and models. *Economic Systems Research,* 11:3**,** 277-300.

MELITZ, M. J. (2003) The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica,* 71:6**,** 1695-1725.

MESNARD, L. (2011) Negatives in symmetric input–output tables: the impossible quest for the Holy Grail. *The Annals of Regional Science,* 46:2**,** 427-454.

MILLER, R. E. & BLAIR, P. D. (2009) *Input-output analysis: foundations and extensions*, Cambridge University Press.

MOSES, L. N. (1955) The stability of interregional trading patterns and input-output analysis. *The American Economic Review,* 45:5**,** 803-826.

NATIONS, U. (1993) *A System of National Accounts.* Brussels, Luxembourg, New York, Paris, Washington D. C.

NITSCH, V. (2000) It’s not right, but it’s okay: On the measurement of intra-and international trade distances. *Manuscript, Bankgesellschaft Berlin*.

NORCLIFFE, G. (1983) Using location quotients to estimate the economic base and trade flows. *Regional Studies,* 17:3**,** 161-168.

O’BRIEN, R. (2007) A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Quality & Quantity,* 41:5**,** 673-690.

OOSTERHAVEN, J. (1981) *Interregional input-output analysis and Dutch regional policy problems.* Aldershot, Gower Publishing Company Limited.

OOSTERHAVEN, J. (1984) A family of square and rectangular interregional input-output tables and models. *Regional Science and Urban Economics,* 14:4**,** 565-582.

OOSTERHAVEN, J. & KNAAP, T. (2003) *Spatial economic impacts of transport infrastructure investments*. In PEARMAN, A., MACKIE, P., NELLTHORP, J. & GIORGI, L. (Eds.) *Transport Projects, Programmes and Policies: Evaluation Needs and Capabilities.* Aldershot, Ashgate.

OOSTERHAVEN, J., PIEK, G. & STELDER, D. (1986) Theory and practice of updating regional versus interregional interindustry tables. *Papers in Regional Science,* 59:1**,** 57-72.

OOSTERHAVEN, J., STELDER, D. & INOMATA, S. (2008) Estimating international interindustry linkages: non-survey simulations of the Asian-Pacific economy. *Economic Systems Research,* 20:4**,** 395-414.

PIISPALA, J. (1999) Constructing Regional Supply and Use Tables in Finland. *Statistics*.

POLENSKE, K. R. (1980) US multiregional input-output accounts and model.

POLENSKE, K. R. & LEVY, P. F. (1975) *Multiregional Economic Impacts of Energy and Transportation Policies*.

REICH, U.-P., STÄGLIN, R. & STAHMER, C. (1995) *Ein System von Input-Output-Tabellen für die Bundesrepublik Deutschland: Konzeption, Erstellung, Auswertung*, Duncker & Humblot.

ROUND, J. I. (1978) An Interregional Input-Output Approach to the Evaluation of Nonsurvey Methods. *Journal of Regional Science,* 18:2**,** 179-194.

ROUND, J. I. (1983) Nonsurvey techniques: a critical review of the theory and the evidence. *International regional science review,* 8:3**,** 189-212.

SARGENTO, A. L. M. (2009) *Regional input-output tables and models: interregional trade estimation and input-output modelling based on total use rectangular tables*. Faculdade de Economia, Universidade de Coimbra.

SCHAFFER, W. A. & CHU, K. (1969) Nonsurvey techniques for constructing regional interindustry models. *Papers in Regional Science,* 23:1**,** 83-104.

SNICKARS, F. & WEIBULL, J. W. (1977) A minimum information principle: theory and practice. *Regional Science and Urban Economics,* 7:1**,** 137-168.

SNOWER, D. J. (1990) New methods of updating input–output matrices. *Economic Systems Research,* 2:1**,** 27-37.

STONE, R. B., A (1962) *A Computable Model of Economic Growth.* London, Chapman and Hall.

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.

VAN RIJCKEGHEM, W. (1967) An exact method for determining the technology matrix in a situation with

secondary products. *Review of Economics and Statistics,* 49.

WIEDMANN, T. (2009) A review of recent multi-region input–output models used for consumption-based emission and resource accounting. *Ecological Economics,* 69:2**,** 211-222.

WIEDMANN, T., WOOD, R., MINX, J., LENZEN, M. & HARRIS, R. (2008) *Emissions embedded in UK trade–UK-MRIO model results and error estimates*. International Input–Output Meeting on Managing the Environment.

WILSON, A. G. (1967) A statistical theory of spatial distribution models. *Transportation research,* 1:3**,** 253-269.

WILSON, A. G. (1970) Inter-regional Commodity Flows: Entropy Maximizing Approaches. *Geographical analysis,* 2:3**,** 255-282.

WILSON, A. G. (1971) A family of spatial interaction models, and associated developments. *Environment and Planning,* 3:1**,** 1-32.

YI, I., ITSUBO, N., INABA, A. & MATSUMOTO, K. (2007) Development of the interregional I/O based LCA method considering region-specifics of indirect effects in regional evaluation. *The International Journal of Life Cycle Assessment,* 12:6**,** 353-364.

Table 8 Variable of the cost structure survey

	Variable
General	Enterprise ID
	Industry
	Location
	Number of employees
Turnover	self-produced commodities (incl. installation, maintenance, hired labor)
	commodities for resale
	wholesaling on a fee
	other activities
	Changes in Inventories
	Internally created property, plant and equipment
Intermediate Consumption	Material inputs
	Energy
	Labor leasing
	Installation, Maintenance etc.
	Renting and leasing
	Interest on borrowed capital
	Insurance premiums
	Social security
	Other social costs
	Other cost: services
Primary inputs	Compensation of employees
	Consumption taxes
	Taxes
	Depreciation
	Subsidies on production
	Value-added taxes (VAT)
	Deducible VAT

Source: own calculations IEK-STE 2013

Table 9 Annual report of establishments

	Variables
General	Establishment ID
	Enterprise ID
	Industry
	Location
	Number of employees
	Compensation of employees
Turnover	in Germany
	Abroad
	Total

Source: own calculations IEK-STE 2013

STE-Preprints 2014

01/2014 Linssen, Jochen, Hake, Jürgen-Friedrich: International Hydrogen Research, Development, Demonstration and Deployment Activities.

02/2014 Schlör, Holger, Hake, Jürgen-Friedrich, Fischer, Wolfgang: Integrated Sustainability Assessment of Climate Impact, Land, Energy and Water Use in Germany against the Background of the UN Green Economy Model and Germany's Sustainability Strategy.

03/2014 Stenzel, Peter, Baumann, Manuel, Fleer, Johannes, Zimmermann, Benedikt, Weil, Marcel: Database development and evaluation for techno-economic assessments of electrochemical energy storage systems.

Research Reports 201

01/2014 Többen, Johannes: Construction of Multi-regional Supply-Use Tables: Experiences from Germany's Federal States

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 energy systems analysis. Technical, economic and ecological subsystems which interact with each other often have to be investigated simultaneously. The group Systems Analysis and Technology Evaluation (STE) takes up this challenge focusing on the long-term supply- and demand-side characteristics of energy systems. It follows, in particular, the idea of a holistic, interdisciplinary approach taking an inter-linkage of technical systems with economics, environment and society into account and thus looking at the security of supply, economic efficiency and environmental protection. This triple strategy is oriented here to societal/political guiding principles such as sustainable development. 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 institutions.

Leitung/Head: Prof. Jürgen-Friedrich Hake

ForschungszentrumJülich

Institute of Energyand Climate Research

IEK-STE: Systems Analysis and Technology Evaluation

52428 Jülich

Germany

Tel.: +49-2461-61-6363

Fax: +49-2461-61-2540,

Email: preprint-ste@fz-juelich.de

Web: www.fz-juelich.de/ste