Calculation of a coherent matrix of intermediate consumption in a situation of incomplete product information

Bayerl, Nikolaus

Statistics Austria, Guglgasse 13, 1110 Vienna, Austria

Ph: +43-1-71128-8263, e-mail: nikolaus.bayerl@statistik.gv.at, http://www.statistik.at

Abstract

According to the data transmission programme of the ESA 1995 the EU member states have to compile annual supply and use tables at a level of 64 activities and 64 products. For several parts of the system product information on a more detailed level is available. This paper shows a way to create a coherent matrix of intermediate consumption which fulfils the necessary balancing conditions for the production and product accounts of the SU-framework and which respects the boundaries given by the detailed product information. The method described in this paper utilizes special information like results of primary statistics, annual accounts or other information and integrates it into a matrix of intermediate consumption classified on the CPA-6-digit-level for the product dimension while the activity dimension remains on the standard NACE-64-level. After the description of the method a simple example shows the process of compilation.

The main advantage of this approach is to get a more detailed view on intermediate consumption than provided in the official use tables. On the one hand the obtained information can be used for the annual balancing process. Especially for the balancing at constant prices a very detailed product breakdown is helpful. On the other hand for analytical purposes a transparent transformation of single business data into aggregated data like national accounts is achieved.

Introduction

In the data transmission programme of the ESA regulation No. 1267/2003 the member states are obligated to generate annual supply and use tables in which the application of the main classifications with 64 industries and 64 products is required. Especially for analysis where a very deep insight of the product input structure for several industries is necessary this minimum level of detail for the products can be a too high aggregation.

More disaggregated product information can give a detailed insight in the inter-industrial linkages. Especially in the case where a producer only creates a few products a very concise picture for the technological inputs is given. In the area of price and volume measures in national accounts as much product information as possible is necessary particularly when the product groups to be calculated at constant prices are very inhomogenous on the one hand and the price development differs very strong between each intermediately used product on the other hand. This can lead to a wide range of intermediate consumption at constant prices depending on whether at what aggregation level is calculated. Moreover the assessment of adequacy of price indices for specific transactions can be improved.

For several parts of the SU-framework there exists product information on a more detailed level for the intermediate consumption, in the ideal case of specific enterprises which can deal as the minimum boundary levels of macroeconomic statistics like national accounts. The data sources for these can be manifold and range from basic primary statistics, single enterprise´s annual accounts, industry analysis to different specific information like newspaper articles, interviews or a personal conversation with an industry expert.
But with that information alone it is not possible to calculate intermediate consumption for the total branch where many different data sources are incorporated. There seems to be a trade-

off between a detailed microeconomic product input mix for single enterprises and macroeconomic aggregate levels for the total intermediate consumption of specific acitivities like in the system of national accounts. The problem tightens even more if the detailed product information which is available in the most cases for individual local kind of activities is contradictory to the finally balanced aggregate data of the SU-tables in which a huge variety of data is processed.

A number of studies deal with the problem of completing IO-models under partial information by estimating missing information on the basis of available data (Wood 2010, Ahmed and Preckel 2007). The treatment of balancing under conflicting information is described at great length in Lenzen et al. (2009).

It should be the target of an adjustment technique to make as much use as possible out of the specific product data under the condition that the boundary values given by the SU-balancing process are respected. The way from the detailed producer´s input data based on single information to total industry data and in the end to the finally balanced SU-tables should be followed. By running feedback-loops potentially arising data conflicts can be minimised.

Underlying basic data

The way of combining different data sources on intermediate consumption to reach plausible coefficients in the SU-tables is carried out by a hierarchical methodological approach going out in the first instance from single producer´s data to the total aggregate data.

Given the use-matrix of the SU-tables $\boldsymbol{U}$ in the dimension of product groups i and activities j.

$$\boldsymbol{U} = \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1j} & \cdots & u_{1m} \\ u_{21} & u_{22} & \cdots & u_{2j} & \cdots & u_{2m} \\ \vdots & \ddots & \ddots\vdots & \ddots\vdots & \ddots\vdots & \vdots \\ u_{i1} & u_{i2} & \ddots\vdots & \ddots\vdots & \ddots\vdots & u_{im} \\ \vdots & \ddots\vdots & \ddots\vdots & \ddots\vdots & \ddots\vdots & \vdots \\ u_{n1} & u_{n2} & \cdots & u_{nj} & \cdots & u_{nm} \end{pmatrix}$$

Each line of this matrix contains the upper boundary values for the intermediate consumption of a specific product group per activity. With that information a huge variety of row vectors $\boldsymbol{u_i}$ containing the upper amount of input of a product group for each activity can be set.

$$\boldsymbol{u_i} = (u_{i1} \quad u_{i2} \quad \cdots \quad u_{ij} \quad \cdots \quad u_{im}) \text{ for } i = 1, \dots, n$$

Given a matrix (square or rectangular) $\boldsymbol{B^i}$ of the minimum values in the size of all available detail products within a product group by all activities classified in the SU-tables.[1]
Some of the coefficients contain several producers´ specific intermediate consumption data whereas the rest of the coefficients are set to null. Moreover the number of columns of matrix $\boldsymbol{B^i}$ equals the number of columns of matrix $\boldsymbol{U}$, in other words the number of activities is the same in matrix $\boldsymbol{B^i}$ as well as in matrix $\boldsymbol{U}$.[2]

[1] In the following demonstration the exponents are used for indices and not for mathematical power.
[2] For the sake of simplicity here all matrices for detail products have the same number of rows. Generally this is not given, then it has to be considered that the number of rows depend on the index i.

$$\boldsymbol{B^i} = \begin{pmatrix} b_{11}^i & b_{12}^i & \cdots & b_{1j}^i & \cdots & b_{1m}^i \\ b_{21}^i & \cdots & \cdots & b_{2j}^i & \cdots & b_{2m}^i \\ \vdots & \ddots & \ddots & \ddots\vdots & \ddots\vdots & \vdots \\ b_{k1}^i & b_{k2}^i & \ddots\vdots & b_{kj}^i & \ddots\vdots & b_{km}^i \\ \vdots & \ddots\vdots & \ddots\vdots & \ddots\vdots & \ddots\vdots & \vdots \\ b_{l1}^i & b_{l2}^i & \cdots & b_{lj}^i & \cdots & b_{lm}^i \end{pmatrix} = \{b_{kj}^i\},$$
$$\text{where } b_{kj}^i \leq 0 \text{ for } i = 1, \dots, n \text{ and } j = 1, \dots, m \text{ and } k = 1, \dots, l$$

A coefficient unequal to zero stands for a minimum amount of a product´s input in an activity whereas a coefficient of zero signifies that no specific data is available but this doesn´t mean that the total industry has no input of that specific product.
But even for that matrix a first constraint has to be valid. The total sum of the coefficients has to be below or equal the aggregate intermediate consumption for each product group in the Use-tables and moreover the total sums of the column sums have to be below or equal the aggregate intermediate consumption of a specific product group for each activity in the Use-tables.

$$\sum_{k=1}^{l}\sum_{j=1}^{m} b_{kj}^i \leq \sum_{m=1}^{m} u_{ij} \text{ for } i = 1, \dots, n, \sum_{k=1}^{l} b_{kj}^i \leq u_{ij} \text{ for } i = 1, \dots, n \text{ and } j = 1, \dots, m$$

Because of the fact that each coefficient of matrix $\boldsymbol{B^i}$ specifies each line of matrix $\boldsymbol{U}$ the column sums of matrix $\boldsymbol{B^i}$ can be seen as the greatest lower boundaries of intermediate consumption per activity.

Given the product vector $\boldsymbol{p^i}$ of the total sums of intermediate consumption per detailed product.

$$\boldsymbol{p^i} = \begin{pmatrix} p_1^i \\ p_2^i \\ \vdots \\ p_q^i \end{pmatrix} \text{ where } \sum_{j=1}^{m} b_{rj}^i \leq p_r^i \text{ for } r = 1, \dots, \min(q, l)$$

In the ideal case the number of rows of vector $\boldsymbol{p^i}$ equals the number of rows of matrix $\boldsymbol{B^i}$, but this condition is non-essential since by calculating a difference coefficient d^i for vector $\boldsymbol{p^i}$ the necessary constraint that the column sum of vector $\boldsymbol{p^i}$ equals the row sum of vector $\boldsymbol{u_i}$ can be established. For the next steps the number of rows of matrix $\boldsymbol{B^i}$ is manually equalized to the number of rows of the vector $\boldsymbol{p^i}$ by vertical concatenating of matrix $\boldsymbol{B^i}$ with a row vector which contains only zeros in the first step. At the beginning of the adjustment this zero-line acts as a placeholder but it is filled with concrete amounts in the following steps. A calculated amount in the obtained line of difference coefficients per activity stands for unknown quantities of product details in intermediate consumption within a product group.

$$\sum_{s=1}^{q} p_s^i + d^i = \sum_{j=1}^{m} u_{ij}$$

In a situation with perfectly reliable information the margins of an unknown target matrix is available. A generalized adjustment technique for introducing non-reliable margins by using random errors is given in Vazquez and Hewings (2011) because this assumption can be unrealistic. Here the margins for the intermediate consumption per product group for an activity is fixed by the reconciled SU-dataset.

With that information a first matrix of known and unknown data including the necessary boundary values per activity and per product can be set. In the following illustration it is assumed that the number of rows of matrix $\boldsymbol{B}^i$ is equal to the number of rows of vector $\boldsymbol{p}^i$, so in that case $q=l$.

b_{11}^i	…	b_{1j}^i	…	b_{1m}^i	p_1^i
…	…	…	…	…	…
b_{k1}^i	…	b_{kj}^i	…	b_{km}^i	p_k^i
…	…	…	…	…	…
b_{l1}^i	…	b_{lj}^i	…	b_{lm}^i	p_l^i
$d_{l+1,1}$	…	$d_{l+1,j}$	…	$d_{l+1,m}$	d^i
u_{i1}	…	u_{ij}	…	u_{im}	

To sum up the matrix to be adjusted contains the minimum values given in matrix $\boldsymbol{B}^i$, the boundary values given in the vectors $\boldsymbol{u}_i$ and $\boldsymbol{p}^i$ plus the difference coefficient d^i and the difference vector $\boldsymbol{d}$ which has only zeros at the beginning of the adjustment.

Hierarchical adjustment method

The adjustment follows a hierarchic way from the scaling of the most disaggregate dataset to the next most aggregate one. In the first step of the hierarchical matrix adjustment the minimum values for the intermediate consumption are fixed by subtracting the column sums of matrix $\boldsymbol{B}^i$ from the coefficients of the use vector $\boldsymbol{u}_i$ on the one hand and by subtracting the row sums of matrix $\boldsymbol{B}^i$ from the coefficients of the product vector $\boldsymbol{p}^i$. The residual boundary values are now given in the vectors $\overline{\boldsymbol{u}}_i$ and $\overline{\boldsymbol{p}}^i$

$$\overline{\boldsymbol{u}}_{\boldsymbol{i}} = (\bar{u}_{i1} \quad \bar{u}_{i2} \quad \cdots \quad \bar{u}_{ij} \quad \cdots \quad \bar{u}_{im}) = \{\bar{u}_{ij}\} \text{ with } u_{ij} - \sum_{k=1}^{l} b_{kj}^i \quad \text{for } i = 1, \dots, n \text{ and } j = 1, \dots, m$$

and

$$\overline{\boldsymbol{p}}^{\boldsymbol{i}} = \begin{pmatrix} \overline{p}_1^i \\ \overline{p}_2^i \\ \vdots \\ \overline{p}_q^i \end{pmatrix} = \{\overline{p}_k^i\} \text{ with } \overline{p}_k^i = p_k^i - \sum_{j=1}^{m} b_{kj}^i \quad \text{for } i = 1, \dots, n \text{ and } k = 1, \dots, q$$

The remaining residual boundary values are now incorporated into a matrix by combining a dummy matrix in a RAS-technique.
Given the dummy matrix $\boldsymbol{D}^i$ which has the same size as matrix $\boldsymbol{B}^i$ plus the row vector of the difference line $\boldsymbol{d}$.

$$\boldsymbol{D}^{\boldsymbol{i}} = \begin{pmatrix} d_{11}^i & d_{12}^i & \cdots & d_{1j}^i & \cdots & d_{1m}^i \\ d_{21}^i & \cdots & \cdots & d_{2j}^i & \cdots & d_{2m}^i \\ \vdots & \ddots & \ddots & \ddots\vdots & \ddots\vdots & \vdots \\ d_{k1}^i & d_{k2}^i & \ddots\vdots & \ddots\vdots & \ddots\vdots & d_{km}^i \\ \vdots & \ddots\vdots & \ddots\vdots & \ddots\vdots & \ddots\vdots & \vdots \\ d_{l1}^i & d_{l2}^i & \cdots & d_{lj}^i & \ddots\vdots & d_{lm}^i \\ d_{l+1,1} & d_{l+1,2} & \cdots & d_{l+1,j} & \cdots & d_{l+1,m} \end{pmatrix} = \{d_{kj}^i\} \text{ where } d_{kj}^i = \begin{cases} 1, & \text{when used} \\ 0, & \text{when not used} \end{cases}$$

The dummy matrix $\boldsymbol{D^i}$ only constitutes of coefficients which are either zero or one. A coefficient equal 1stands for a constellation when it is known that a specific product is used by an activity in the production process but the exact amount of that input is unknown. In such a case an additional input of that product should be added up to the minimum values given in matrix $\boldsymbol{B^i}$. In the other case the input should remain 0 after the RAS-adjustments. Setting of either 1 or 0 should not only be done for the product lines but for the difference line too.

There are four main cases for a coefficient of 0. First of all if the boundary value for the product is still reached after the subtraction of the minimum values in $\boldsymbol{B^i}$ from the use vector $\boldsymbol{u_i}$ or the product vector $\boldsymbol{p^i}$ no additional inputs should be added. In the second case if an activity doesn't use a specific product within a product group this can be seen by a zero in the use-matrix. For these two cases the dummy coefficient in $\boldsymbol{D^i}$ should be set equal to 0 but basically it would not change the result of the RAS-iterations if they were set to 1.
In the next two cases it plays a major role to set the dummy values to 0. If the information is available that a specific product is only used in some discrete activities the dummy variables should only be there set to 1 whereas for the others to 0. In the other case when the specific input mix for several industries is known only these products should set to 1 for these activities and the rest to 0.
The advantage of utilizing this special information is that not only for the known industries the possible estimation errors are reduced because of targeted allocating of inputs to the correctly using activities.

The RAS-procedure is an iterative scaling method of biproportional adjustment of rows and columns whereby here the non-negative matrix $\boldsymbol{D^i}$ is pre-multiplied with a diagonal matrix of row factors of corrections and post-multiplied with a diagonal matrix of column factors of corrections. The correction factors come from the known vectors $\boldsymbol{\overline{u}_i}$ and $\boldsymbol{\overline{p}^i}$ including the difference coefficient, this means only the margins of an unknown target matrix $\boldsymbol{D^t}$ are available. The iterations v are continued until the column sums and row sums converge to given target vectors this means the solution is a matrix that diverges least with respect to the prior and is consistent with the aggregate information observed for the target.

Generally speaking it is the principle to find a target matrix $\boldsymbol{T}$ which fulfills the following target equation:

$$\boldsymbol{T} = \lim_{v\to\infty} {}_v\overline{P}^i{}_{2(v-1)}D^i{}_v\overline{U} \text{ and } \sum_{i=1}^{n} T_{ij} = \overline{p}_i^i \text{ and } \sum_{j=1}^{m} T_{ij} = \overline{u}_j$$

Given ${}_{2(v-1)}\boldsymbol{D^i}$ as the matrix which arises from the iteration $v\text{-}1$ then the first iteration matrix with the diagonal matrix ${}_v\boldsymbol{\overline{P}^i}$ is calculated as given.

$${}_{2(v-1)+1}\boldsymbol{D^i} = {}_v\overline{P}^i{}_{2(v-1)}D^i, \text{where } {}_v\overline{P}_i^i = \left\{ {}_v\overline{p}_i \right\} = \frac{T_{i.}}{{}_{2(v-1)}D_{i.}^i}$$

After that step the row sum of the resulting matrix ${}_{2(v-1)+1}\boldsymbol{D^i}$ is equal to the row sums given in the target vector $\boldsymbol{\overline{p}^i}$ including the difference coefficient.
In the next step the matrix ${}_{2v}\boldsymbol{D^i}$ is calculated by using the diagonalised matrix ${}_v\boldsymbol{\overline{U}}$.

$${}_{2v}\boldsymbol{D}^i = {}_{2(v-1)+1}D^i{}_v\overline{U}, \text{where } {}_v\overline{U}_j = \left\{ {}_v\overline{u}_j \right\} = \frac{T_{.j}}{{}_{2(v-1)+1}D_{.j}^i}$$

Starting from the dummy matrix $\boldsymbol{D^i}$ as ${}_0\boldsymbol{D^i}$ it is checked after the end of this first iteration if the target equation is still reached. Otherwise the next iteration starts with the resulting matrix

$_{2v}\boldsymbol{D}^{i}$. This is done until a matrix with the same dimensions, properties and margins as the target matrix ***T*** is reached. The last row of this matrix stands for not available information on product details for the intermediate consumption per industry which is needed to fulfill the boundary values given in the SU-tables.

In the next step the matrix $\boldsymbol{Z}^{i}$ of intermediate consumption in the dimension product details plus difference line by industries is calculated by adding the matrix $\boldsymbol{B}^{i}$ of the minimum values plus the zero line to the target matrix ***T***. This matrix still fulfills the balancing conditions for the production and product accounts of the SU-framework and is consistent with the boundaries given by the detailed product information.

$$\boldsymbol{Z}^{i} = \begin{pmatrix} b_{11}^{i} & b_{12}^{i} & \cdots & b_{1j}^{i} & \cdots & b_{1m}^{i} \\ b_{21}^{i} & \cdots & \cdots & b_{21}^{i} & \cdots & b_{2m}^{i} \\ \vdots & \ddots & \ddots & \ddots\vdots & \ddots\vdots & \vdots \\ b_{k1}^{i} & b_{k2}^{i} & \ddots\vdots & b_{kj}^{i} & \ddots\vdots & b_{km}^{i} \\ \vdots & \ddots\vdots & \ddots\vdots & \ddots\vdots & \ddots\vdots & \vdots \\ b_{l1}^{i} & b_{l1}^{i} & \cdots & b_{lj}^{i} & \ddots\vdots & b_{lm}^{i} \\ 0 & 0 & \cdots & 0 & \cdots & 0 \end{pmatrix} + \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1j} & \cdots & t_{1m} \\ t_{21} & \cdots & \cdots & t_{2j} & \cdots & t_{2m} \\ \vdots & \ddots & \ddots & \ddots\vdots & \ddots\vdots & \vdots \\ t_{i1} & t_{i2} & \ddots\vdots & \ddots\vdots & \ddots\vdots & t_{1m} \\ \vdots & \ddots\vdots & \ddots\vdots & \ddots\vdots & \ddots\vdots & \vdots \\ t_{n1} & t_{n2} & \cdots & t_{nj} & \ddots\vdots & t_{nm} \\ t_{n+1,1} & t_{n+1,2} & \cdots & t_{n+1,j} & \cdots & t_{n+1,m} \end{pmatrix} = \{z_{ij}^{i}\}$$

$$\text{where} \sum_{j=1}^{m} z_{ij}^{i} = u_j \;\forall\, j \text{ and } \sum_{i=1}^{n} z_{ij}^{i} = p_i \;\forall\, i$$

If the resulting matrix $\boldsymbol{Z}^{i}$ does not seem to be plausible there are two ways of manual corrections. First of all the minimum values of matrix $\boldsymbol{B}^{i}$ can be increased manually. For that way the constraints for the row and column sums have to be accounted for. A second way of manual correction works by defining taking and giving industries. If the product input of a specific activity seems to be too low this industry can be identified as a taking industry and it is clear that there has to be at least one giving industry, in the most cases there will be more than one. In the first step the taking industry gets an amount of a specific product from the giving industry. After that all the constraints for the column sums for the taking as well as for the giving industries are not fulfilled any more whereas the boundaries for the row sums remain unaffected. So in the next step the taking industry pays for the receipt of the product by transferring other own products to the giving activities.

In the end of these steps a matrix of intermediate consumption is available which is on the one the hand coherent with the boundary values of the SU-balancing process as well as the detailed product information and on the other hand a plausible and circumstantial description of the use of products in the production process.

Possible Feedback loops

With that process a deeper insight in an underlying product technology is given. The process described above is not only a formal way of adjusting a matrix to given boundary values but an instrument for the plausibilization of data by running feedback loops. In that case the plausibility check can be done either for the SU-balancing process by a check column by column or for statistics on the intermediate consumption with detailed product information by a check line by line.

After merging different data sources on intermediate consumption discrepancies in the form of technological impossibilities can occur. If a specific column sum, the SU-value, is below the minimum value this is a sign of a too low allocation of a good for an activity in the process of structuring total intermediate consumption to aggregate product groups.

The constellation if a row sum is below the minimum value can be interpreted as a calculation error for a statistic on total intermediate consumption classified by detailed products. Moreover if the minimum value matrix contains products which are not given in the aggregate statistics there are data gaps which have to be filled.

A numerical example

On the basis of a small numerical example the method of adjusting a detailed matrix on intermediate consumption going out from specific information given at different aggregate levels as described above is illustrated. For the sake of simplicity there is given an economy with only the three activities coffeehouse, yoghurt producer and sweet producer on the one hand and the four input products coffee beans, milk, sugar and water.

For these activities some detailed information on intermediate consumption is available. An owner of a coffeehouse tells that an amount of 100 of coffee beans, 200 of milk and 50 of sugar is used in the production process. According to the annual report of another coffeehouse 400 of coffee beans, 600 of milk, 150 of sugar and 300 of water is used intermediately. A newspaper article about a specific yoghurt producer reports that an amount of 250 of milk is used.
Besides that detail information there are given the boundary values of the SU-tables in the dimension product group by activity and of the vector on total intermediate consumption classified by product details.
On the basis of that information the following matrix system can be compiled containing the minimum value matrix, the zero line for the aggregate good food products and the boundary values according to the SU-balancing process as well as to the detailed product statistics. Because of the fact that the detailed product statistic in that example does not match the boundary values given by the reconciled SU-tables a difference coefficient in the amount of 3000 has to be calculated.

product group	detailed products	coffeehouse	yoghurt producer	sweet producer	product detail boundary value
food products	coffee beans	500	0	0	1500
	milk	800	250	0	4500
	sugar	200	0	0	2500
	water	300	0	0	3500
	aggregate good	0	0	0	3000
	SU boundary value	7000	6000	2000	

The given boundary values are now reduced by the minimum values to receive the target vectors for the RAS-adjustment. The remaining input for the activity coffeehouse is now 5200, for the yoghurt producer 5750 and for the sweet producer still 2000. The input of coffee beans is now 1000, of milk 3450, of sugar 2300 and of water 3200.

Now an activity expert tells that coffeehouses typically use coffee beans and milk for their products but the exact amount of the total activity is unknown. Moreover the expert knows that besides of the two coffeehouses where minimum values are available no other coffeehouse uses sugar and water. This means that no additional sugar and water should be allocated to the activity coffeehouse.
A report on product technology contains the information that for producing yoghurt no coffee beans are needed but milk, sugar and water. And beyond this information it can be seen from that report that no other products as sugar and water are used in the activity sweet producer. Because of that data the dummy matrix looks like that.

product group	detailed products	coffeehouse	yoghurt producer	sweet producer	product detail boundary value
food products	coffee beans	1	0	0	1000
	milk	1	1	0	3450
	sugar	0	1	1	2300
	water	0	1	1	3200
	aggregate good	1	1	0	3000
	SU boundary value	5200	5750	2000	

In that example 33 iterations are necessary to get a result that fulfils the constraints for the column and row sums.

product group	detailed products	coffeehouse	yoghurt producer	sweet producer	product detail boundary value
food products	coffee beans	1000	0	0	1000
	milk	2246,5	1203,5	0	3450
	sugar	0	1463,6	836,4	2300
	water	0	2036,4	1163,6	3200
	aggregate good	1953,5	1046,5	0	3000
	SU boundary value	5200	5750	2000	

After adding the minimum values the preliminary matrix on total intermediate consumption has the following structure. This matrix still fulfills the constraints of the product and production accounts of the aggregate SU-tables.

product group	detailed products	coffeehouse	yoghurt producer	sweet producer	product detail boundary value
food products	coffee beans	1500	0	0	1500
	milk	3046,5	1453,5	0	4500
	sugar	200	1463,6	836,4	2500
	water	300	2036,4	1163,6	3500
	aggregate good	1953,5	1046,5	0	3000
	SU boundary value	7000	6000	2000	

Now the national accountant and the activity expert say that the input structure of the sweet producer seems not to be plausible because there should be 200 more input of sugar. One possibility would be to set this specific value into the minimum value matrix and run the process again. Another way is to define the sweet producer as a taking activity and the yoghurt producer as a giving activity for the input of the product sugar. So the input of sugar for the sweet producer is increased and for the yoghurt producer reduced by 200. But the sweet producer has to remunerate the yoghurt producer. Because of the fact that the sweet producer only has the two inputs sugar and water this can only be carried out by a decrease of water for the sweet producer and an increase for the yoghurt producer in the same amount.

In the end the final matrix of intermediate consumption is obtained.

product group	detailed products	coffeehouse	yoghurt producer	sweet producer	product detail boundary value
food products	coffee beans	1500	0	0	1500
	milk	3046,5	1453,5	0	4500
	sugar	200	1263,6	1036,4	2500
	water	300	2236,4	963,6	3500
	aggregate good	1953,5	1046,5	0	3000
	SU boundary value	7000	6000	2000	

In this example possible feedback loops can be demonstrated. A technological impossibility can occur for example if there are some big coffeehouses which tell that they all together have an input of coffee beans in the amount of 8000. Of course this value must be checked for reliability before any further steps are set. For the SU-balancing process it must be tested if the column sums are not exceeded by the minimum values. In that case it is clear that the SU-boundary value of 7000 is too low, so the input of the product group food products in the activity coffeehouse has to be increased in the process of structuring this activity´s total intermediate consumption.
The check for the statistics on the product details has to be arranged row by row. In this example the minimum value exceeds the boundary value by far which is a sign of a serious mistake in the statistics on product details because the boundary value has to be at least as high as the minimum value. Moreover if the owner of a coffeehouse says that in addition to the products in the example cream is used there is a data gap in the statistics on the product details where this input good is not covered.

Conclusions

The process described in this paper shows a practical way of compiling a matrix utilizing different data sources at different levels of product aggregations which fulfills the condition of inner coherence. The more information is available the better will be the result of intermediate consumption and the underlying product technology. With this method the way from single business data to aggregate data like national accounts can be traced.

References

Ahmed, A.; Preckel, P.; *A comparison of RAS and entropy methods in updating IO tables*, Economics Association Annual Meeting, Portland, Oregan, USA, 2007

Wood, R.; *Generic methods for estimating input-output models under partial information*, 18th International Conference on Input-Output Techniques, Sydney, Australia, 2010

Vazquez, E., Hewings J.D.; *Matrix adjustment with non reliable margins: a Generalized Cross Entropy approach*, 19th International Conference on Input-Output Techniques, Alexandria, USA, 2011

Lenzen, M., B. Gallego and Wood, R.; *Matrix balancing under conflicting information*, Economic Systems Research 21(1), 2009

Robinson, S., Cattaneo, A. and El-Said, M.; *Updating and estimating a social accounting matrix using cross entropy methods*, Economic Systems Research 13 (1), 2001