26th International Input-Output Conference in Juiz de Fora (Brazil), 25th to 29th of June 2018

About Upper and Lower Bounds of Spatial Aggregation Effects.

Dietrich W. Koeppen

University of Duisburg-Essen, Department of Economics

D r a f t

Abstract

The upper and lower limits of interregional feedbacks in input-output models have played a role in theoretical and practical discussions of these multiplier effects. A parallel can be drawn to the analysis of biases which result if the aggregation level of interregional input-output systems is changed.

What can be said about the over or underestimation of intra and interregional trade multipliers, comparing results of a given macro model M* with those on a deeper level M with a more extended spatial differentiation. The question has a practical background if, for example, interregional input-output investigations have to work, typically because of empirical restrictions, with aggregates like the rest of the economy.

The paper presents an approach to determine the positive or negative direction of possible distortions but also verifies their upper and lower limits, thus giving hints on the reliability of the original results in case of unknown relations on a deeper level. The estimation procedure uses less information as possible, serving eventually as a first step into further more cost intensive research.

The method is based on the comparison of input or output coefficients and final demand or primary input proportions using power series. All effects and limits can be observed for whole regions as well as single industries or industry groups in each region.

The partition of spatial aggregates into sub-regions is flexible. It can be shown under which conditions the aggregation effects reach a maximum or a minimum, if they matter or can be neglected and in which cases they disappear, confirming the original results.

From a methodological viewpoint, the concept makes clear in which way spatial but also sectoral aggregation effects depend on the distribution, especially the concentration of supplies and deliveries together with final demand relations. This extends the usual assumption that the biases are caused by the heterogeneity of units.

Introduction.

Formally, there are no principle differences between a sectoral or a spatial aggregation of I-O systems. Industries as well as regions with intermediate transactions and relations to exogenous parts like final demand and primary inputs, can be grouped at a basic level after various patterns to form aggregated units at a higher level, the operation connects a macro system and a micro system by an aggregation function. In both cases, analytical consequences are unavoidable: if the model results, especially the multipliers, gained by a micro system are summed up they differ normally from the aggregated macro results. A central topic of aggregation theory is to define and to explain this effect which is usually denoted as aggregation bias.

The theory of sectoral aggregation in single I-O models has been focused during a broad discussion on the conditions for a consistent or acceptable transformation of micro systems into macro systems (a survey is given in Kymn [1990], see also Olson [2001], Dietzenbacher [1991] presents a special approach). This search has to do, among others, with an interest to concentrate complex I-O structures at one side and the intention not to change the model results at the other side.

Under a spatial viewpoint, this interest has shifted. Although the possibilities for regional I-O studies have been widely improved by surveys, non-survey or hybrid methods (see for example Lahr [2001]) the data situation is still difficult, especially to construct complete interregional models including all transactions between a greater number of spatial units in a sufficient sectoral differentiation.

Therefore its of some interest to know about the relevance of spatial aggregation effects, especially whether larger I-O systems are necessary to verify relevant feedbacks and spillovers, or if the distortions which may occur if the number of regions is reduced are small enough to be neglected.

One can try to answer this question by starting from a given full developed interregional I-O model and then observe the effects after aggregating the regions in different ways (see for example extended experiments with Japanese data and also with chance generated data in Miller and Blair [1988, 2oo9]). The biases were usually found to be small, and as a consequence the necessity of large interregional models covering all trade flows between the regions has been doubted [1]. But at least in case of experiments with real systems one may argue that the results are just reflecting a special situation.

The paper presents another approach, beginning with the general form of an interregional macro model and comparing this with a disaggregated micro model [2]. The method will be demonstrated by investigating a given I-O system of one or more regions with transactions to the rest of the economy known only as aggregates. It can be shown in which way hidden aggregation effects, responsible for over and under estimations of the original feedbacks and spillovers at the macro level, depend on relations of deliveries and supplies between assumed sub-regions measured as input coefficients, without knowing these transactions precisely. Thus, the coefficients which are necessary to calculate interdependencies must not be estimated exactly but its sufficient to observe their relative weights, especially the concentration on certain sub-regions.

Not only the direction of such distortions can be derived but also their upper and lower limits [3]. They make it possible to calculate intervals for the searched aggregation effects and so to say something about their relevance in a concrete case, the main intention of this paper.

Technically, the estimation procedure is based on a power series development, allowing to distinguish different steps of influences between final demand parts and regional production as overall effects as well as distortions at the industry level.

As a kind of methodological by-product, some final hints are given connecting the aggregation issue with the linear and circular structure of I-O systems.

2. Basic Relations.

To analyze how spatial aggregation changes the results of interregional I-O models, a standard demand driven quantity version [4] shall be observed on two aggregation levels, a macro model M* with

(1.1) $\mathbf{X}^* = (\mathbf{I} - \mathbf{A}^*)\,\mathbf{f}$

and a micro system on a deeper level M

(1.2) $\mathbf{X} = (\mathbf{I} - \mathbf{A})\,\mathbf{f}$

The m by m block matrix **A*** contains the input matrices of m regions at the macro level, the system **A** those of n regions of the micro system, in each case for s industries. X* and X, f* and f include vectors of dimension s for regional output and final demand.

Both aggregation levels are connected by a system **G** with m unity and zero matrices grouping the sub-regions in (1.2) to the aggregated regions in (1.1)

(1.3) $$\mathbf{G} = \begin{bmatrix} \mathbf{I}\ldots\mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}\ldots\mathbf{I} & \mathbf{0} & \mathbf{0} \\ & & & \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}\ldots\mathbf{I} \end{bmatrix}$$

together with a corresponding weight system

(1.4) $$\mathbf{W} = \begin{bmatrix} \mathbf{W}_{11} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_{12} & \mathbf{0} & \mathbf{0} \\ . & . & & . \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{W}_{mm} \end{bmatrix}$$

weighting the sub-regions by their total output shares.

The macro system then can be transformed into the micro system (and in reverse) by

(1.5) $\quad \mathbf{A^*} = \mathbf{GAW'} \quad$ and $\quad \mathbf{A} = \mathbf{W'A^*G}$

This leads to the difference

(1.6) $\quad \mathbf{D} = [\,(\mathbf{I} - \mathbf{GAW'})^{-1}\mathbf{G} - \mathbf{G}(\mathrm{I} - \mathbf{A})^{-1}]\ \mathbf{f}$

as the usual definition of aggregation effects.

The equation can be used to observe these effects directly, calculating all differences between the aggregation levels, especially the over and the underestimation of feedbacks and spillover effects for concrete interregional systems [4]. But it can also be used for another perspective: what can be said about D in case of a given macro system M* if the information about the structure of M is restricted, which are the conditions of over- or underestimations concerning the original model results and are there upper and lower limits of these virtual aggregation effects.

3. An estimation approach for spatial aggregation effects.

3.1 A bi-regional case.

The analysis shall be concentrated first on a bi-regional constellation where the principles of the estimation procedure can be shown clearly but the derivations become less complex (for extensions see section 3.4 below).[3]

The frame is defined by a region 1, denoted as observation region, which is embedded in a national economy. The intraregional I-O data are given but also the interregional relations with the rest of the economy. Such constellation is of some practical meaning if, at the one side, the incentives and especially the data and research resources to construct large economy covering regional I-O projects are insufficient, but at the other side, there is enough initiative to build a single region model, from a metropolitan study up to large regions, its then plausible to try at least an extension where the connections with the rest of the economy are observed.

The results derived from such a bi-regional model give at least hints about feedbacks and spillovers between the observation region and the rest of the economy [4], but normally they are affected by aggregation errors. These biases, their positive or negative direction, their upper and lower limits and hence their importance, shall be investigated with help of the relations (1.1) - (1.6).

The matrix system A* then includes 4 coefficient matrices

(2.1) $$\mathbf{A}^* \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{1R} \\ \mathbf{A}_{R1} & \mathbf{A}_{RR} \end{bmatrix}$$

To demonstrate the estimation procedure, the rest of the economy R shall be split into two sub-regions R_2 and R_3 (for more than 2 sub-regions see also section 3.4). The lower level system **A** is then

(2.2) $$\underline{\mathbf{A}} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \mathbf{A}_{13} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{A}_{23} \\ \mathbf{A}_{31} & \mathbf{A}_{32} & \mathbf{A}_{33} \end{bmatrix}$$

The sub-regions can be grouped and weighted by

(2.3) $$\mathbf{G} = \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{I} \end{bmatrix}$$ (2.4) $$\mathbf{W} = \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_2 & \mathbf{W}_3 \end{bmatrix}$$

Aggregated regional output and final demand are

(2.5) $$\mathbf{GX} = \mathbf{X}_1 \quad \mathbf{X}_2 + \mathbf{X}_3$$ (2.6) $$\mathbf{GF} = \begin{bmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 + \mathbf{f}_3 \end{bmatrix}$$

The choice of the sub-regions is open, depending on observation interests and available data.

Following a usual approximation the aggregation biases in (1.6) can be expressed by a power series

(2.7) $$\mathbf{D} = (\mathbf{GAW'G} - \mathbf{GA})\,\mathbf{f} + (\mathbf{GA^2W'G} - \mathbf{GA^2})\,\mathbf{f} \; . \; . \; . + (\mathbf{GA^zW'G} - \mathbf{GA^z})\,\mathbf{f}$$

The vectors in **D** measure all differences between macro and micro results from the viewpoint of the original level and can be differentiated after their exogenous origins:

(2.8) $$\mathbf{D} = \begin{bmatrix} \mathbf{D}_{11} + \mathbf{D}_{1R} \\ \mathbf{D}_{R1} + \mathbf{D}_{RR} \end{bmatrix}$$

This splitting of the aggregation bias is substantial because the effects can be analyzed separately:

- $\mathbf{D}_{11}$ concerns feedbacks [5] in case of exogenous changes of final demand F_1 in region 1 which have influence on this region´s output,

- $\mathbf{D}_{R1}$ connects the output in the observation region 1 with the final demand in the rest of the economy R and shows possible distortions of these spillover effects,

- $\mathbf{D}_{1R}$ measures aggregation effects in R caused by final demand changes in region 1,

- $\mathbf{D}_{RR}$ represents biases in R caused by the final demand in R.

From the viewpoint of the observation region, these effects are of different relevance. Of special interest are apparently $\mathbf{D}_{11}$ and $\mathbf{D}_{R1}$ which say something about the reliability of the aggregated multipliers concerning the observation region itself.

But it can be investigated also by $\mathbf{D}_{1R}$ and $\mathbf{D}_{RR}$ in which way the rest of the economy is affected by aggregation effects.

3.2 Aggregation effects of first and second order.

Since Theil (1957), the theory has worked successfully with the power series development (2.7) to investigate sectoral aggregation effects, mainly to find conditions for a perfect, unbiased aggregation of industries. The same approach can be used for a spatial aggregation and the more general case of unknown coefficients where the direction and the limits shall be found and where a zero bias is only a special case.

Using a power series has certain advantages. The interdependencies between the production units, in this case regions, are made visible step by step based on transactions and coefficients which can be estimated directly (instead of observing multipliers).

Usually, these steps are not connected with time, they do not reflect a dynamic process in the proper sense. But they give the impression of a sequence beginning with a certain impulse and then continuing with decreasing intensity. So the first terms are of special importance also for possible aggregation errors.[7]

Using (2.8) together with the transformation systems (2.4) and (2.5), the aggregation effects of first order are

$$(2.8) \quad \mathbf{D}_{(1)} = \begin{bmatrix} \mathbf{A}_{11} + (\mathbf{A}_{12}\,\mathbf{W}_2 + \mathbf{A}_{13}\,\mathbf{W}_2)\,\mathbf{f}_1 \\ (\mathbf{A}_{21}+\mathbf{A}_{31})\,\mathbf{f}_1 + [(\mathbf{A}_{22}+\mathbf{A}_{32})\mathbf{W}_2 + (\mathbf{A}_{23} + \mathbf{A}_{33})]\,(\mathbf{f}_2 + \mathbf{f}_3) \end{bmatrix} - \begin{bmatrix} \mathbf{A}_{11}\mathbf{f}_1 + \mathbf{A}_{12}\mathbf{f}_2 + \mathbf{A}_{13}\mathbf{f}_3 \\ (\mathbf{A}_{21}+\mathbf{A}_{31})\,\mathbf{f}_1 + (\mathbf{A}_{22}+\mathbf{A}_{23})\,\mathbf{f}_2 + (\mathbf{A}_{32}+\mathbf{A}_3)\mathbf{f}_3 \end{bmatrix}$$

Apparently, there are no aggregation effects of first order in two cases,

(2.9) $\mathbf{D}_{11(1)} = 0$ and $\mathbf{D}_{R1(1)} = 0$

concerning the relations between final demand in region 1 and the output in this observation region and also to the output in the rest of the economy.

The other conditions in (2.8) are

(2.10) $\mathbf{D}_{1R(1)} = (\mathbf{A}_{12}\mathbf{W}_2 + \mathbf{A}_{13}\mathbf{W}_3)(\mathbf{f}_2 + \mathbf{f}_3) - (\mathbf{A}_{12}\,\mathbf{f}_2 + \mathbf{A}_{13}\,\mathbf{f}_3)$ and

(2.11) $\mathbf{D}_{RR(1)} = [(\mathbf{A}_{22}+\mathbf{A}_{32})\,\mathbf{W}_2 + (\mathbf{A}_{23}+\mathbf{A}_{33})\,\mathbf{W}_3] - [(\mathbf{A}_{22}+\mathbf{A}_{23})\,\mathbf{f}_2 + (\mathbf{A}_{32}+\mathbf{A}_{33})\,\mathbf{f}_3]$

The absence of aggregation effect in case of (2.9) is known[7] and has to do with an observation region 1 which is not disaggregated. But it should be checked whether this holds only for first order biases and so it seems necessary to derive at least in these cases also the second order effects which are

(2.13) $\mathbf{D}_{11(2)} = [\mathbf{A}_{11}\mathbf{A}_{11} + (\mathbf{A}_{12}\mathbf{W}_2+\mathbf{A}_{13}\mathbf{W}_3)(\mathbf{A}_{21}+\mathbf{A}_{31})]\,\mathbf{f}_1 - (\mathbf{A}_{11}\mathbf{A}_{11} + \mathbf{A}_{12}\mathbf{A}_{21} + \mathbf{A}_{13}\mathbf{A}_{31})\,\mathbf{f}_1$

(2.14 $\mathbf{D}_{R1(2} = [(\mathbf{A}_{21}+\mathbf{A}_{31})\mathbf{A}_{11}+[(\mathbf{A}_{22}+\mathbf{A}_{32})\mathbf{W}_2 + (\mathbf{A}_{23}+\mathbf{A}_{33})]\,\mathbf{f}_1 - (\mathbf{A}_{21}\mathbf{A}_{11} + \mathbf{A}_{22}\mathbf{A}_{21} + \mathbf{A}_{23}\mathbf{A}_{31})\,\mathbf{f}_1$
$- (\mathbf{A}_{31}\mathbf{A}_{11} + \mathbf{A}_{32}\mathbf{A}_{21} + \mathbf{A}_{33}\mathbf{A}_{31})\,\mathbf{f}_1$

As assumed, the coefficients are unknown (except for the observation region) but there are possibilities to say something about **D** if their relations can be observed. Note also that the weights W_2 and W_3 can be chosen after the analyst´s interests.

3.3 Over- and underestimation, upper and lower limits: total aggregation effects at the regional level

From the viewpoint of a single region embedded in a national economy, especially the impacts on the total and the sectoral output of this observation region are of interest which come from exogenous variations of final demand within the region itself, among others because of competence reasons. This directs the attention at first on the aggregation effects D_{11} in (2.6) and (2.13).

That first order effects are zero means, a positive or negative choc by the final demand produced in the observation region has a certain influence on the intraregional output independent from any des-aggregation in the rest of the economy. But this first impact induces in a second step a feedback which changes the output also in this rest region.

So the question remains if its possible to estimate (a) the direction of these distortions and (b) their relevance if only the aggregated data are available. To answer this, the equation (2.8) can be transformed into

(3.1) $\mathbf{D}_{11(2)} = (\mathbf{A}_{12} - \mathbf{A}_{13})\,\mathbf{A}_{31}\mathbf{W}_2 + (\mathbf{A}_{13} - \mathbf{A}_{12})\,\mathbf{A}_{21}\mathbf{W}_3$

This makes clear how the type of feedbacks which depends on the influence of final demand on the observation region is determined by relations between the input coefficients of this observation region and those of both sub-regions.

To concentrate the analysis first on the total aggregation effects at the regional level, all s by s coefficients $a_{i,j}$ at the industry level of each spatial unit shall be summed up

(3.1.1) $\sum a_{n,m\,;\,i,j} = a^*_{n,m}$ $\quad i, j = 1 \ldots s \quad n, m = 1 \ldots 3$

The following conditions then can be derived directly from (3.1)

(3.1.2) $d^*_{11(2)} > 0$ if $a^*_{12} > a^*_{13}$ together with $a^*_{31}w^*_2 > a^*_{21}w^*_3$

which means a *positive* bias concerning the second order effects of the original model M*. The feedbacks between the observation region 1 and the rest of the economy representing impacts of the final demand in the observation region on this region are *overestimated*. Apparently, such a distortion occurs if the outgoing deliveries from region 1 to one of the sub-regions, for example region 2, are greater than those to the other region 3 (measured by input coefficients) and at the same time the supplies of the observation region are concentrated on region 3 (note that the regions can be exchanged).

The opposite case, a *negative* bias stands for an *underestimation*, resulting if both the deliveries and the supplies are mainly directed to the same sub-region 2 or 3.

(3.1.3) $d^*_{11(2)} < 0$ if $a^*_{12} > a^*_{13}$ and at the same time $a^*_{31}w_2 < a^*_{21}w_3$

As a third possibility, the aggregation effects can *disappear* if the relevant coefficients are equal:

(3.1.4) $d^*_{11(2)} = 0$ if either $a^*_{12} = a^*_{13}$ or $a^*_{31}w_2 = a^*_{21}w_3$

This special case, with the same model results at both aggregation levels represents a kind of a perfect disaggregation, in analogy to an unbiased sectoral aggregation. Note that only the coefficients on the output or the input side, have to be balanced.

Thus, the existence of positive, negative or zero biases depends apparently on the distribution of outgoing and incoming transactions between the observation region and the sub-regions. As larger the difference between the coefficients, as larger the aggregation effects.

These interregional flows can be concentrated, at least theoretically, on the output and the input side completely on one of the sub-regions, either on the same or on the opposite one. In this case, the aggregation effects reach their upper and lower limits.

If in (3.1) all deliveries go to region 2 and all supplies come from region 3, then these effects reach a maximum:

(3.1.5) $d^*_{11(2)}max = a^*_{12}w_2\, a^*_{31}$ if $a^*_{13} = 0$ and $a^*_{21} = 0$

The sub-regions can be exchanged

(3.1.6) $d^*_{11(2)}max = a^*_{13}w_3a^*_{21}$ if $a^*_{12} = 0$ and $a^*_{31} = 0$

Because in each case one of the coefficients is zero one can also write

(3.1.7) $d^*_{11(2)}max = a^*_{1R}\, a^*_{R1}$

Which means that this upper limit can directly be calculated with help of the given data at the macro level and one gets also an indication of its importance: the aggregation effect is apparently large if the transactions from the observation region to the rest of the economy have weight in this aggregate and the flows from the rest into region 1 are remarkable too.[8]

The lower limit is reached if the observation region is connected by transactions with only one of the sub-regions, the same on the input and the output side.

(3.1.7) $d^*_{11(2)}min = - a^*_{12}w_3a^*_{21}$ if $a^*_{13} = 0$ together with $a^*_{31} = 0$ and

(3.1.8) $d^*_{11(2)}min = - a^*_{13}w_2a^*_{31}$ if $a^*_{12} = 0$ and also $a^*_{21} = 0$

Again the original coefficients can be used

(3.1.9) $d^*_{11(2)}min = a^*_{11} + a^*_{1R}\, a^*_{R1}$

So the bias cannot smaller than zero.

Certainly, the extreme constellations in (3.1.5) to (3.1.8) with their strict concentration may be rather seldom, but the upper and lower bounds show at least the possible margins and can serve as hints on the importance of biases at the regional level. It makes also sense to use these limits because they must not be estimated, they are given directly by the original data, the aggregated transactions to the rest of the economy.[9]

A next group of aggregation effects which are especially important for the observation region concerns the impacts of final demand changes $\mathbf{f_R}$ in the rest of the economy on the output of region 1. Possible biases are measured by $\mathbf{D_{R1}}$ in (2.4).

From the definition (2.6) follows after transformations

(4.1) $\mathbf{D_{R1(1)}} = (\mathbf{A_{12}} – \mathbf{A_{13}})\, \mathbf{f_3}\, \mathbf{W_2} + (\mathbf{A_{13}} – \mathbf{A_{12}})\, \mathbf{f_2}\, \mathbf{W_3}$

This time, differences between the macro and the micro level appear already as first order effects. The total effects at the regional level are positive

(4.1.1) $d^*_{R1(1)} > 0$ if $a^*_{12} > a^*_{13}$ together with $f_3w_2 > f_2w_3$ (or if $a^*_{12} < a^*_{13}$ and $f_3w_2 < f_2w_3$)

They will be negative

(4.1.2) $d^*_{R1(1)} < 0$ if $a^*_{12} > a^*_{13}$ and $f_3w_2 < f_2w_3$ (or $a^*_{12} < a^*_{13}$ with $f_3w_2 > f_2w_3$)

and they disappear in case of

(4.1.3) $d^*_{R1(1)} = 0$ if either $a^*_{12} = a^*_{13}$ or $f_2w_3 = f_3w_2$

These conditions show again that the spatial aggregation causes distortions which depend on the different transactions between the observation region and the sub-regions, but in this case also on the distribution of final demand between these sub-regions. Other than the input coefficients, final demand parts are exogenous variables which can be used for predictions. This plays a role if the upper and lower bounds of the aggregation effects $D_{R1(1)}$ are searched.

The distortions can reach a maximum with

(4.1.4) $d^*_{R1(1)}max = a^*_{12}f_3w_2$ or $a^*_{13}f_2w_3$

and have a lower limit in

(4.1.5) $d^*_{R1(1)}min = - a^*_{12}f_2W_3$ or $- a^*_{13}f_3w_2$

In these cases, a change of final demand in only one of the sub-regions is responsible for the aggregation effects together with a concentration of transactions from the observation region to the other sub-region. The first condition depends on assumptions, the second corresponds to the situation above in (3.1.3) and (3.1.4).

Less important from the observation region´s view point but also of interest as possible distortions at the macro level are the aggregation effects D_{1R} and D_{RR} concerning the production of the rest of the economy R. They can be analyzed in the same way.

Influences on R coming from the final demand in region 1 can be estimated by

(4.2) $\mathbf{D_{1R(1)}} = [(\mathbf{A_{22}} + \mathbf{A_{32}}) - (\mathbf{A_{23}} + \mathbf{A_{33}})]\, \mathbf{A_{31}W_2} + [(\mathbf{A_{23}} + \mathbf{A_{33}}) - (\mathbf{A_{22}} + \mathbf{A_{32}})]\, \mathbf{A_{21}W_3}$

and the intraregional effects where the final demand in the rest of the economy has an impact on the output of this aggregate are

(4.3) $\mathbf{D_{RR(1)}} = [(\mathbf{A_{22}} + \mathbf{A_{32}}) - (\mathbf{A_{23}} + \mathbf{A_{33}})]\, \mathbf{f_3W_2} + [(\mathbf{A_{23}} + \mathbf{A_{33}}) - (\mathbf{A_{22}} + \mathbf{A_{32}})]\, \mathbf{f_2W_3}$

In both cases, the transactions $A_{23} + A_{33} = A_{2R}$ and $A_{22} + A_{32} = A_{3R}$ between the sub-regions 2 and 3 and the rest of the economy, measured by input coefficients, have to be compared, together with the input relations of region 1 or with the final demand proportions between f_2 and f_3.

All conditions for positive and negative aggregation effects can be verified in analogy to (3.2) and (4.1). Finally, its deciding if differences between the sub-regions exist concerning their relations to the observation region, to the rest of the economy and to final demand. The upper and lower limits depend again on the concentration of the relevant transactions.

3.4 Aggregation effects of higher order.

The previous derivations have been used to estimate aggregation effects of first and second order. They are certainly of special importance if real situations and model results shall be observed. But it seems necessary at least of analytical reasons to ask for effects of higher order. Exemplary, they shall be derived for the feedbacks connecting exogenous changes of final demand in the observation region to this region´s output.

From the third term in (2.7)

(5.1) $\mathbf{D}_{(3)} = (\mathbf{GA^3W'G} - \mathbf{GA^3})\,\mathbf{f}$

one obtains for $\mathbf{D}_{11(3)}$

(5.2) $\mathbf{D}_{11(3)} = \mathbf{A}^3_{11} + (A_{12}W_2 + \mathbf{A}_{13}\mathbf{W}_3)\,[(\mathbf{A}_{21} + \mathbf{A}_{31})\,\mathbf{A}_{11} + [\mathbf{A}_{11} + (\mathbf{A}_{22}+ \mathbf{A}_{32})\mathbf{W}_2 + (\mathbf{A}_{32} + \mathbf{A}_{33})\mathbf{W}_3]]$

$- \,[\,(\mathbf{A}_{11}\mathbf{A}_{12}+\mathbf{A}_{12}A_{21}+\mathbf{A}_{13}\mathbf{A}_{31})\,\mathbf{A}_{11} + (\mathbf{A}_{11}\mathbf{A}_{21}+\mathbf{A}_{12}A_{22}+\mathbf{A}_{13}\mathbf{A}_{32})\,\mathbf{A}_{21} + (\mathbf{A}_{13}A_{11}+\mathbf{A}_{12}\mathbf{A}_{23}+\mathbf{A}_{13}\mathbf{A}_{32})\mathbf{A}_{31}\,]$

The relations between the coefficients become more complex because exogenous impulses, after their entrance into the system, spread out via deliveries and supplies over the whole net of transactions. Not only $\mathbf{A}_{12}$ and $\mathbf{A}_{13}$ at the output side of the observation region together with $\mathbf{A}_{21}$ and $\mathbf{A}_{31}$ at the input side play a role for aggregation effects but also $\mathbf{A}_{22} + \mathbf{A}_{32}$ and $\mathbf{A}_{32} + \mathbf{A}_{33}$ within the rest of the economy. But the feedbacks at the macro level as well as the micro level decrease, so at least the absolute value of the aggregation bias becomes also smaller.

Whether the effect changes from under- to overestimation or in reverse is not to be expected because the additional intraregional influences within the aggregate go into the same direction as the other ones.

Its presumably of interest to observe the total effects overlooking all following steps of the power series. The simplest way is then to solve both models M* and M in the usual way (1.6) and (1.7). This becomes easier in case of upper and lower bounds which can be calculated by the given macro data. Also equal coefficients in case of zero biases can be found directly. For other constellations, the relevant coefficients could eventually gained by assuming close or wide intervals [10].

3.4 Aggregation effects at the sectoral level

To observe the influence of a spatial aggregation in detail at the sectors level makes sense among others because the overall effects normally balance stronger differences between industries. One may also focus the analysis on special cases, for example key industries or those with important interregional dependencies.

Assuming such a selection, the equations found above can be noted as conditions for sector 1 s_1 , the industry of interest, and sector 2 s_2 including all other industries. The procedure shall be demonstrated for sectoral influences of final demand on the production in region 1 $\mathbf{D_{11}}$ an .

The condition (3.1) $(\mathbf{A_{12}} - \mathbf{A_{13}})\,\mathbf{A_{31}W_2} + (\mathbf{A_{13}} - \mathbf{A_{12}})\,\mathbf{A_{21}W_3}$ then becomes

(6.1) $\mathbf{D_{11(2)}}$ =

$$\begin{bmatrix} (a_{12,11} - a_{13,11})a_{31,11}w_2 + (a_{12,12} - a_{13,12})a_{31,21}w_2 & (a_{12,11} - a_{13,11})a_{31,12}w_2 + (a_{12,12} - a_{13,12})a_{31,22}w_2 \\ (a_{12,21} - a_{13,21})a_{31,11}w_2 + (a_{12,21} - a_{13,22})a_{31,21}w_2 & (a_{12,21} - a_{13,11})a_{31,12}w_2 + (a_{12,22} - a_{13,22})a_{31,22}w_2 \end{bmatrix}$$

$$+ \begin{bmatrix} (a_{13,11} - a_{12,11})a_{21,11}w_3 + (a_{13,12} - a_{12,12})a_{21,21}w_3 & (a_{13,11} - a_{12,11})a_{21,12}w_3 + (a_{13,12} - a_{12,12})a_{21,22}w_3 \\ (a_{13,21} - a_{12,21})a_{21,11}w_3 + (a_{13,22} - a_{12,22})a_{21,21}w_3 & (a_{13,21} - a_{12,21})a_{21,12}w_3 + (a_{13,22} - a_{12,22})a_{21,22}w_3 \end{bmatrix}$$

Now its possible to check all aggregation effects concerning the relation between both sectors in region 1 and those sectors in the rest of the economy resulting from a final demand change in this observation region. The difference $d_{11\,(2),\,11}$ for example measures the second order bias in case of transactions from the observed industry 1 in region 1 to the same industry in the aggregated region.

(6.1.1) $d_{11(2),\,11} = (a_{11,\,12} - a_{11,\,13})\, a_{11,31}\, w_2 + (a_{11,13} - a_{11,12})\, a_{11,21}\, w_3$

Analog to ( ) to ( ) the direction of this sectoral effect becomes clear:

(6.1.2) $d_{11(2),11} > 0$ if $a_{11,12} > a_{11,13}$ together with $a_{11,31}\, w_2 > a_{11,21}\, w_3$

(6.1.3) $d_{11(2),11} < 0$ if $a_{11,12} > a_{11,13}$ with $a_{11,31}\, w_2 < a_{11,21}\, w_3$ (the subscripts can be exchanged)

There is no bias in case of

(6.1.4) $d_{11(2),11} = 0$ if $a_{11,12} = a_{11,12}$ or $a_{11,21}\, w_2 = a_{11,31}\, w_3$

So the concentration of sectoral deliveries and supplies to and from one of the sub-regions is responsible for an over or an underestimation at the original macro level and the aggregation effect disappears if the transactions are balanced.

It is of course possible to use the conditions (4.2) and (4.3) to detect aggregation effects concerning spillovers between region 1 and the rest of the economy and feedbacks within for the aggregate at the industry level.

With help of

(4.2.1) $d_{1R,\,i\,j\,(1)} = [(a_{22,\,i\,j} + a_{32,\,i\,j}) - (a_{23,\,i\,j} + a_{33,\,i\,j})]a_{31,\,i\,j}\,w_2 + [(a_{23,\,i\,j} + a_{33,\,i\,j}) - (a_{22,\,i\,j} + a_{32\,i\,j})\,a_{21,i\,j}\,w_3$

and

(4.2.2) $d_{RR,\,i\,j\,(1)} = [(a_{22,\,i\,j} + a_{32,\,i\,j}) - (a_{23,\,I\,j} + a_{33,\,i\,j})]\,f_{3,\,I}\,w_2 + [(a_{23,\,i\,j} + a_{33,\,i\,j}) - (a_{22,\,i\,j} + a_{32,\,I\,j})\,f_{2,\,i}\,w_3$

All s by s distortions concerning deliveries from region 1 to the rest of the economy and intra-regional transactions in this aggregated region for relations between any industry i and j can be observed in detail

3.4 About extensions.

Out of several possibilities to extend the analysis of spatial aggregation effects based on the approach used above, the following shall be noted briefly.

In case of multiregional models where the interregional links are not or only partly known but the relations to the rest of the economy for the single regions are given as aggregates, the aggregation effects with over and under estimations, upper and lower bounds could be observed separately for each region. The given multiregional data would probably improve the estimations.

One could also try to get a deeper look into a spatial aggregate by distinguishing more than two sub-regions. That means a comparison of at least 3 estimated coefficients at the input and at the output side of an observation region and also of the corresponding final demand parts.

The procedure can be based for n sub-regions on these equations:

(7.1) $\mathbf{A_{1R}} = \mathbf{A_{12}W_2} + \mathbf{A_{13}W_3} + \ldots + \mathbf{A_{1n}W_n}$

(7.2) $\mathbf{A_{R1}} = \mathbf{A_{21}} + \mathbf{A_{31}} + \ldots + \mathbf{A_{n1}}$

(7.3) $\mathbf{A_{RR}} = (\mathbf{A_{22}W_2} + \mathbf{A_{32}W_3} + \ldots + \mathbf{A_{n2}W_2}) + \ldots + (\mathbf{A_{2n}W_n} + \ldots + \mathbf{A_{nn}W_n})$

(7.4) $\mathbf{f_R} = \mathbf{f_2} + \mathbf{f_3} + \ldots + \mathbf{f_n}$

To start the estimation for example for D_{11} which connects final demand in the observation region 1 with the output of this region, its useful to fix the order of the sub-regions either at the output side in (6.1) or at the input side in (6.2) and then check the other ranking:

a negative aggregation effect (of second order) results. There is a positive bias in case of

(7.1.2) $\mathbf{A_{21}} > \mathbf{A_{31}} > \ldots > \mathbf{A_{n1}}$ and at the same time $\mathbf{A_{12}W_2} < \mathbf{A_{13}W_3} < \ldots < \mathbf{A_{1n}W_n}$

Its of course possible that such parallel orders cannot be achieved if for example

(7.1.3) $A_{12}W_2 < A_{13}W_3$ but $A_{13}W_3 > A_{14}W_4$

In this case, additional assumptions about the absolute values of the coefficients are necessary.

The other types of aggregation effects at the regional and at the sectoral level and also their upper and lower limits can be derived in a similar way using the conditions which were found above.

To observe the relevant rankings is certainly more difficult than in case of only two sub-regions. In general, finding out the aggregation effects of interest will depend again on searching at least stronger or weaker tendencies for a concentration or a balance of the transactions between one observation region (or several) and the sub-regions and within the rest aggregate.

As shown above, upper and lower limits are connected with an extreme concentration which means setting zero all coefficients in the rows (7.1.) to (7.1.) and similar ones, except one at the output and one at the input side.

4. Summary.

The method which was sketched above to investigate the structure of spatial aggregates in rest-of constellations, especially with regard to unknown aggregation effects, is based on a core of given I-O data for one or more observation regions which are confronted with the rest of the economy or another spatial aggregate and tries to need, depending on the way how the aggregate shall be partitioned, a minimum of additional estimations.

The procedure starts plausibly with the definition of sub-regions. Their choice is flexible reflecting, for example, interests in special interdependencies between observation regions and sub-regions. If enough data are available different pattern could be tested.

In the next step, informations or at least hypotheses are needed about intra and interregional transactions measured in terms of input coefficients together with assumptions about final demand shares .

The method works with a separate estimation of deliveries from the observation region to the sub-regions and supplies of this region and transactions within the aggregate together with the observation of final demand shares. The deciding task is to find out in which way the flows are concentrated or balanced. A role for relevant assumptions could play, among others, the regions´ peripheral or central position and but also their production structure which is eventually concentrated on special products or services.

The upper and lower bounds which can also be derived have a special meaning, concerning cases of a full concentration of the relevant transactions on one (or a group) of the sub-regions they can be obtained directly from the macro data and so provide intervals which do not have to be estimated but are given. An extreme spatial concentration may not seem to be very realistic but in cases of a strong tendency the effects could come close to these limits, at least on the industry level.

Whether the effects are remarkable or stay less important in a real constellation depends on this distribution but a frame is given already by the weights of the interregional trade flows for the observation regions and the rest of the economy, and this has to do also with their relative size. In case of a large rest aggregate, the weight of the incoming supplies tend to be rather small and in reverse. A main role plays of course in which degree the observation regions depend on outgoing and incoming flows.

The estimation approach has definitively a preparing character. Once a kind of basic structure concerning the interregional flows, especially with regard to concentrations or balances, can be made visible, more precise results, eventually focused on special questions, could be obtained by further investigations.

Some final methodological remarks. All derivations are based on a standard input-output model, that means under special assumptions about production (or output) functions, quantity and price reactions. So, their validity is bound to these assumptions [9]. But staying in the logic of this model centered around dependencies and interdependencies, one can find some hints why aggregation effects occur.

Observing a macro system where an observation part of one or several regions is connected with a spatial aggregate by deliveries and supplies, these output and input flows suggest first a circular context. Such circularity may be confirmed by analyzing the input and output structure within the aggregate, but its also possible that the interdependencies are restricted or disappear. This is especially the case if the observation regions and the sub-regions are connected by a chain where the transaction follow one main direction establishing dependencies rather than interdependencies. An aggregation then can hide this context leading to an overestimation of feedbacks and spillovers.

If at the other hand circular relations exist only for a part of a system, for example between one observation region and one sub-region, the aggregation tends to underestimate the effects.

Real sectoral as well as regional input-output systems are usually characterized by a mix of linear and circular relations, but it seems of some interest, especially in a spatial sub- or transnational context, to investigate the dominant tendencies. The sensibility for aggregation effects is only one and presumably a rather secondary aspect but provides at least a better look on results which were obtained at the original level.

Notes

[1] The spatial aggregation effects found by Miller and Blair (1988) especially for feedbacks in one region (non aggregated) against between one and seven aggregated regions if final demand changes region 1 were below one percent. They concluded that, for example, a "two-region" model for California against the rest of the U.S. would be sufficient. And in general (p 163) for "total regional outputs associated with new final demands the construction of a many-region interregional input-output model appears unnecessary."

[2] Discussing sectoral aggregation, this perspective was sometimes denoted as micro-bias approach (see Kymn [1990] p 74)

[3] Upper bounds on the sizes of interregional feedbacks were especially investigated by Miller (1986) and Guccione, Gillen, Blair and Miller (1988) using, among others, vector norms.

[4] The approach can be adapted for other basic quantity and prize I-O models, using output instead of input coefficients and working with primary inputs instead of final demand as exogenous parts. For this duality see for example (Olsen 1993, 2001).

[5] As usual, the interregional multiplier effects resulting from impacts on the regional production caused by final demand changes in the same region are denoted as feedback effects, the influences coming from other regions as spillovers. The whole issue of aggregation effects is closely related to these basic elements of interregional I-O systems, their existence and their meaning. The discussion is reassembled for example in Round (2001).

[6] There are many attempts to endogenize the external relations of a single region by constructing bi-regional I-O systems where this region is connected with the rest of the economy. One reason is often the existence of national I-O data which provide the completing of such a model, assumed the trade flows between both parts are available or can be derived.

[7] Miller and Blair (1981), among others, refer to this by their *Theorem II* : "If some sectors are not aggregated and the new final demands occur only in unaggregated sectors, the first order aggregation bias will vanish." This is of course true also for total regions.

[8] The following numerical examples may give an impression of the possible size of the aggregation bias in case of $d_{11(2)}$.

Assumed the macro data for an observation region R_1 and the rest of the economy R_R expressed by input coefficients are

$A^*_{11} = 0,3$ $A^*_{1R} = 0,15$ $A^*_{R1} = 0,14$

(possible output data and transactions could be for example $x_1 = 1000$ $x_R = 1000$ with intraregional flows in R_1 of 300, deliveries from R_1 to R_R = 150 and supplies of R_1 from R_R = 140)

The second order feedback effect is then $A_{11}A_{11} + A^*_{1R} A^*_{R1} = 0,111$ (which means that a change of final demand of 1 in the observation region R_1 results in change of output in this region of 0,111).

To test this result, a division of the aggregate R_R into the sub-regions R_2 and R_3 shall be assumed with weights $W_2 = W_3 = O,5$. The upper limit is reached if the observation region delivers all its interregional transactions to one of the sub-regions 2 or 3 and gets all its supplies from the other on. So $A_{12} = 0,3$ $A_{13} = 0$ $A_{21} = 0$ $A_{31} = 0,14$ and the feedback effect at the micro level is then $A_{11}A_{11} + A_{12}A_{21} + A_{13}A_{31} = A_{11}A_{11} = 0,09$

Thus, there is an absolute difference between both levels of $d_{11(2)} = + 0,021$. The relative difference can be calculated if one relates this difference to the original result 0,111 0,021/0,111 = 0,189 . the aggregated model overestimates this feedback with nearly 19 per cent, a rather remarkable bias.

Because a full concentration on one of the sub-regions is not very realistic one can assume a weaker one. If the observation region has transactions with R_2 and R_3 but still with an overweight for on sub-region, for example with the following distribution $A_{12} = 0,26$ $A_{13} = 0,o4$ $A_{21} = 0,01$ $A_{31} = 0,13$ then the feedback at the micro level is 0,0978 which means a difference of 0,0132 and the relative aggregation effect remains still 0,119 (nearly 12 p.c.)

Of course its interesting to observe the opposite constellation where the observation regions concentrates deliveries and supplies on the same sub-region. Assuming the coefficients from above, this time with $A_{13} = 0$ and $A_{31} = 0$, a lower limit will be reached with d = - 0,0232 , this means an underestimation of 0,209 21p.c., and in case of a something weaker concentration still – 15,13 p.c.

These positive and negative aggregation effects are clearly more important than those reported by Miller and Blair (1981, 2009). Although hypothetical examples should be observed with some caution, one could draw the consequence to look for the concentration of transactions between an observation region and the parts of a spatial aggregate also for real interregional constellations and not to trust too much on the assumed irrelevance of unknown biases.

[9] A comparison of a classical I-O with a CGE approach in a bi-regional model for Scotland and the rest of the UK (but without regarding aggregation effects) is given by McGregor, Swales, Yin (1999)

References.

Dietzenbacher, E. (1992) : Aggregation in Multisector Models: Using the Perron Vector. *Economic Systems Research, Vol 4, 1*

Howe, E.C. and Stabler, J.C. (1989): Canada Divided: the optimal Division of an Economy into Regions. *Journal of Regional Science, Vol 29,2*

Guccione, A., Gillen, W.J., Blair, P.D., Miller, R.E. (1988): Interregional Feedbacks in I-O Models: The Least Upper Bounds. *Journal of Regional Science, Vol 28*

Kymn, K.O. (1990): Aggregation in Input-Output Models: A Comprehensive Review. *Economic Systems Research, Vol. 2*

Lahr, M. L. (2001): A Strategy for Producing Hybrid Regional Input-Output Tables. In *Input-Output Analysis – Frontiers and Extensions.* Lahr, M. L. and Dietzenbacher, E. (eds.)

McGregor, P. G., Swales, J. K., Yin, Y. P. (!999) in Equlibrium Interregional Models of the National Economy: A Requiem for Interregional Input-Output? In *Understanding and Interpreting Economic Structure.* Sonis, M. and Hewings, G. J. D., Madden, M., Kimura, Y. (eds.)

Miller, R. E. and Blair, P. D. (!981): Spatial Aggregation in Interregional Input-Output Models. *Papers of the Regional Science Association, 48*

Miller, R. E. (1986): Upper Bounds on the Size of Interregional Feedbacks in I-O models. *Journal of Regional Science, Vol. 26*

Olson, J. A. (1993): Aggregation in I-O Models: Prices and Quantities, *Economic Systems Research V.5*

Olson, J. A. (2001): Perfect Aggregation of Industries in I-O Models. In *Input-Output Analysis – Frontiers and Extensions.* Lahr, M. L. and Dietzenbacher, E. (eds.)

Round, J. I. (2oo1): Feedback Effects in Interregional I-O Models: What Have We Learned. In *Input-Output Analysis – Frontiers and Extensions.* Lahr,M. L. and Dietzenbacher, E. (eds.)