UNIVERSITÀ POLITECNICA DELLE MARCHE

DIPARTIMENTO DI ECONOMIA

ASSESSING THE BEHAVIOUR OF NON-SURVEY METHODS OF CONSTRUCTING REGIONAL INPUT-OUTPUT TABLES THROUGH A MONTE CARLO SIMULATION

ANDREA BONFIGLIO AND FRANCESCO CHELLI

QUADERNO DI RICERCA n. 293

Giugno 2007

Assessing the Behaviour of Non-Survey Methods of Constructing Regional Input-Output Tables through a Monte Carlo Simulation

Andrea Bonfiglio and Francesco Chelli*

Abstract

The paper aims to analyse the tendency of a battery of non-survey techniques of constructing regional I-O tables to over-(under-)estimate impact. The behaviour of the regionalization methods is assessed relatively to the techniques analysed. For this aim, a Monte Carlo simulation has been carried out. Then, a multidimensional scaling procedure has been applied to search for a common and repeated structure of differences among the methods and to give an immediate picture of possible implications, in terms of impact direction, coming from the choice of a given regionalisation method rather than another. Afterwards, the results have been compared to those obtained by applying the same procedure to 2000 I-O tables, which have been mechanically constructed for the 20 Italian regions.

The results indicate that the choice of the regionalization method is crucial in estimating multipliers. According to the chosen method, the extent of multipliers could be considerably bigger or lower. This can have serious repercussions in terms of policy choices and, therefore, policy makers and I-O analysts should be aware of it.

In addition, the results have confirmed a tendency of the methods to over-(under)-estimate impact both statistically and empirically. However, they have also shown that sectoral aggregation can reverse this tendency. Finally, from an economic point of view, it turned out that the most recent Flegg et al. Location Quotient (Flegg et al., 1995; Flegg and Webber, 1997) is the best to represent regional economies.

Keywords: non-survey techniques, impact analysis, Monte Carlo simulation, multidimensional scaling procedure, regional policy

J.E.L. Classification: C15, C67, R15

---

* Dr. Andrea Bonfiglio – Department of Economics, Faculty of Economics "G. Fuà", Polytechnic University of the Marche, Piazzale Martelli, 8, 60121 Ancona (Italy). Tel.: +39 071 2207116, Fax.: +39 071 2207102, e-mail: a.bonfiglio@univpm.it. Prof. Francesco Chelli – Department of Economics, Faculty of Economics "G. Fuà", Polytechnic University of the Marche, Piazzale Martelli, 8, 60121 Ancona (Italy). Tel. 071-2207054, Fax. 071-2207102, e-mail: f.chelli@univpm.it. We are grateful to an anonymous referee for remarks and suggestions on a previous version of the paper.

Contents

1 Introduction

An input-output table, which often represents the basis for developing SAMs, requires the knowledge of all flows of goods and services among intermediate and final sectors expressed in a disaggregated form and related to a given time period. This implies the collection of a great volume of information, which, at a sub-national level, is difficult to make ready for use. For this reason, alternative approaches for deriving regional input-output tables have been developed over time. Three main approaches can be identified as "survey", "non-survey" and "hybrid" approaches.

The survey-based approach attempts to identify the elements of a transactions table from a collection of primary data by surveys of industries and final consumers, concerning both sales and purchases. Non-survey techniques derive the elements of a transactions table from other (usually national) tables by various modification techniques. Finally, the hybrid approach combines non-survey techniques with superior data, which are obtained from experts, surveys and other reliable sources (primary or secondary). It is therefore a compromise between the survey and non-survey approaches in order to reduce costs associated with survey-based models and to reach satisfactory levels of reliability, overcoming the main problems related to non-survey methods.

Currently, hybrid methods (Jensen et al., 1979; Greenstreet, 1989; West, 1990; Midmore, 1991; Lahr, 1993; Jackson, 1998; Madsen and Jensen-Butler, 1999; Lahr, 2001) and ready-made models (Brucker et al., 1987; Jensen, 1987; Round, 1987; Treyz et al., 1992; Lindall and Olson, 1998) are the most used by I-O analysts.

However, this does not mean at all that non-survey methods are being no more employed. On the contrary, in spite of several criticisms to mechanical procedures (Round, 1987), non-survey techniques have been revalued thanks to the diffusion and a widespread use of ready-made models, which are fundamentally based on these types of techniques. Therefore, many models of impact prediction and evaluation are definitively based on non-survey techniques.

In addition, non-survey techniques are widely employed within hybrid procedures to derive a first estimate of coefficients which are then adjusted to conform to exogenous information (Jensen et al., 1979; Lahr, 2001). Adjustments are generally made by optimization techniques which minimise differences between indirectly estimated coefficients and final coefficients, under constraints represented by accounting identities and/or exogenous information. Therefore, the resulting I-O table and multipliers are likely to be affected by the choice of the non-survey method used.

A widespread use of indirect techniques of constructing regional tables, which represent the basis for impact and, in general, regional analysis, raises the problem of verifying their tendency to over-(under-)estimate impact. This is because policy choices based on the use of Input-Output analysis or CGE models can produce effects which differ considerably from initial expectations engendered by the regionalisation method adopted. The main risk is that an exaggerated multiplier can push policy makers to unjustified public concessions to private industry, which can consist of land acquisition, new infrastructure, job training programs, subsidized loans, tax facilities and so on. The extent of public concessions is determined by the bargaining between policy makers and industry and during the phase of negotiations, it is likely that those who benefit most from the project tend to inflate multipliers.

In literature, there are several empirical studies which demonstrate that conventional regionalization techniques yield substantial overestimates of I-O coefficients (for instance, Morrison and Smith, 1974; Harrigan et al., 1980; Stevens et al., 1989) as well as there are studies which demonstrate that there are alternative techniques which allow the solution to the problem of underestimating regional imports and overestimating regional multipliers (Flegg and Webber, 1997; Tohmo, 2004). However, most studies base their conclusions on given regional cases and we cannot be sure that these results are valid in general. Moreover, they focus on one or some regionalization methods and thus it is not possible to obtain a more complete and updated picture regarding the behaviour of the various methods which have been proposed over time.

Hence, the need for an approach, which collects a sufficiently wide number of methods, including the most recent ones, and is able to verify if the results provided separately by empirical studies can be generalised, is unquestionable.

The aim of the paper is to analyse the tendency of a battery of techniques of constructing regional I-O tables to over-(under-)estimate impact through a both statistical and empirical approach. The behaviour of the regionalization methods is assessed relatively to the techniques analysed. In other words, impacts estimated by regionalization methods are compared to each other and to not to a benchmark impact, i.e. a “true” impact coming from a real I-O table. This is because the objective is not to evaluate whether one method is better than another one but to attempt to provide policy makers and analysts with an immediate picture of possible implications, in terms of impact direction, coming from the choice of a given regionalisation method rather than another.

The paper is articulated as follows. The next section illustrates the regionalisation methods investigated. The third section describes the methodology of analysis adopted. The forth section finally provides some concluding remarks.

2 The methods investigated

The regionalization methods analysed are part of the location quotients family. They are: the Simple Location Quotient (SLQ), the Purchases-only Location Quotient (PLQ), the Cross-Industry Location Quotient (CILQ), the Semilogarithmic Location Quotient (RLQ), the Symmetric Cross Industry Location

Quotient (SCILQ) and four versions of the Flegg et al. Location Quotient (FLQ). The choice of these methods depends on the fact that they are the most debated and those appearing most frequently in literature. In addition, they require few data to be applied and thus they are well suited to be analysed by the methodology adopted in this research.

All the non-survey techniques considered are aimed at estimating regional input coefficients assuming that regional and national technologies are identical. The regional input coefficient is estimated in the following way: $a_{ij}^R = q_{ij} r_{ij}^N$, where r_{ij}^N is the national technical coefficient, q_{ij} represents the percentage of modification of the national coefficient. The interregional import coefficient is usually estimated by the difference between the regional (national) technical coefficient and the regional input coefficient. The non-survey techniques analysed can be considered as different ways of estimating q_{ij}.

For a wider illustration of these methods see Miller and Blair (1985), Flegg et al. (1995), Flegg and Webber (1997), Oude Wansink and Maks (1998). Here, only a brief description will be given.

The SLQ takes the following form:

$$SLQ_i = \frac{X_i^R / X^R}{X_i^N / X^N}$$

where i indexes a given sector, X is output, R and N indicate the region and the nation, respectively. For reasons related to data unavailability at a local level, output data are often replaced with employment or value added data. It is established that $\text{SLQ}_i = 1$ when $\text{SLQ}_i \geq 1$. The SLQ is applied uniformly along the rows of the national technological matrix. The logic behind is that if a regional sector is relatively less important than the same sector at a national level ($\text{SLQ}_i < 1$), the regional sector will not be able to satisfy all local requirements and a part of supply will be imported. In a contrary case ($\text{SLQ}_i \geq 1$), the regional sector will be able to satisfy the local demand completely.

The PLQ is as follows:

$$PLQ_i = \frac{X_i^R / X^{*R}}{X_i^N / X^{*N}}$$

where X^* is output of only those industries that use i as input. The PLQ is applied as the SLQ.

The CILQ takes the following form:

$$CILQ_{ij} = \frac{X_i^R / X_i^N}{X_j^R / X_j^N}$$

This location quotient is applied as the SLQ with the difference that the national matrix is adjusted cell by cell.

The SCILQ is one variant of the traditional CILQ. It is designed to take into consideration the possibility of deriving regional coefficients that exceed the national ones, overcoming the problem of asymmetric adjustments. It takes the following form:

$$SCILQ_{ij} = 2 - \frac{2}{CILQ_{ij} + 1}$$

The logic behind this is the following. If the CILQ equals zero or one, then the SCILQ equals zero or one. If the CILQ goes to infinity, the SCILQ goes to two. In so doing, the regional coefficients do not only take account of the fact that sectors may be less concentrated in a region, but also that sectors may be more concentrated.

The RLQ takes the following form:

$$RLQ_{ij} = \frac{SLQ_i}{\log_2 \left(1 + SLQ_j\right)}$$

According to Round (Round, 1972; 1978), this method incorporates the properties of both the SLQ and CILQ methods. In other words, the RLQ takes account of the importance of the region, of the selling sectors and of the purchasing sectors.

The FLQ is a modification of the RLQ and takes the following form:

$$FLQ_{ij} = \begin{cases} CILQ_{ij} \cdot \lambda^* & \text{for } i \neq j \\ SLQ_i \cdot \lambda^* & \text{for } i = j \end{cases}$$

where $\lambda^* = \left[\log_2 \left(1 + X^R / X^N\right)\right]^{\delta}$, $0 \leq \delta < 1$, $0 \leq \lambda^* \leq 1$. The application of the SLQ along the main diagonal is motivated by the need to eliminate the problem of overestimation of intrasectoral coefficients. This issue will be discussed in par. 4 more extensively. The FLQ is designed to incorporate the properties of the CILQ and the SLQ, eliminating the relevant shortcomings. The use of the FLQ requires identifying the δ parameter. If the value of δ is higher, the adjustment for regional imports will be greater. So, this parameter is inversely related to the size of the region.

3 The methodology used

The methodology employed to analyse the behaviour of the regionalisation methods is based on a Monte Carlo simulation. 100 national 100-sector I-O matrices were generated randomly. For every national matrix, one national and one regional employment vector, simulating the employment structure of the hypothetic nation and region, were derived randomly.[1] We imposed that national employment in every sector could not exceed 1,000 labour units and that employment in every regional sector could not be bigger than the corresponding national level.[2] Using employment vectors, 9 location-quotient-based methods were applied to derive 9 corresponding regional I-O coefficient matrices. The used methods are: the SLQ, the PLQ, the CILQ, the RLQ, the SCILQ and four versions of the FLQ (FLQ with $\delta = 0.1$, $\delta = 0.3$, $\delta = 0.6$, $\delta = 0.9$, respectively). From the I-O coefficient matrices, the relevant Leontief inverse matrices were derived and the relevant sector multipliers were calculated.

In order to explore relationships among regionalization methods, a multivariate statistical procedure is applied: the multidimensional scaling procedure (MSP). A similar approach has been already used in Bonfiglio and Chelli (2004) to analyse the behaviour of non-survey methods in estimating the impact of SAPARD pre-accession instrument in rural regions of Slovenia, Romania and Bulgaria.

MSP is a set of related statistical techniques used for exploring similarities or dissimilarities in data. The objective is to represent a given matrix of similarities or dissimilarities among n objects, in a k-dimensional space, suitable for graphing or 3D visualisation.[3]

To apply MSP, sector multipliers calculated from every regionalization methods were compared to each other by calculating a Euclidean distance for every sector. Then an average of Euclidean distances was calculated until obtaining a squared double-entry matrix, which represents a synthesis of the dissimilarities existing

[1] Obviously, a different meaning such as output or value added can be attributed to the vector.

[2] This limit was chosen arbitrarily. However, any other limit could be imposed but this would not have modified the results.

[3] This procedure attempts to find a structure from a set of distance measures among objects. This operation is carried out assigning observations to specific positions within a reduced conceptual space, in order to make distances among points on the space correspond to specified dissimilarities as much as possible. In this way, it is possible to obtain a representation of least-squares of objects within the space, which mostly helps to understand data in a better way. The procedure was applied using the Software Package SPSS 13 (PROXSCAL procedure).

among the methods in estimating impacts. As many distance matrices were derived as there were randomly generated national I-O matrices. MSP was applied to all the distance matrices organised as piled matrices across columns.

4 Empirical results

Results derived from the Monte Carlo simulation are shown in Table 1 and Figure 1.

Table 1 shows the average value of sector output multipliers calculated for every regionalization method after 100 iterations and the corresponding ranking. According to the table, the method which produces on average the biggest multiplier is the SCILQ, followed by the SLQ, the PLQ, the RLQ, the CILQ, and all the versions of the FLQ starting from the FLQ with a lower value of the parameter δ until the FLQ with the highest value of the parameter. Therefore, in comparison with the other techniques, the SCILQ tends to overestimate multipliers whereas the FLQ tends to underestimate impact. Clearly, overestimation generated by the SCILQ derives from the possibility offered by this method of estimating regional coefficients which can be bigger than the corresponding national coefficients. The remaining methods produce middle results. In particular, we can note that the RLQ produces an average multiplier which is very close to those estimated by the SLQ and the PLQ. Therefore, although the RLQ was presented as a method able to overcome drawbacks related to the SLQ and the CILQ, the results show that it does not make difference roughly using the SLQ and the RLQ in estimating multipliers. In addition, it turns out that the CILQ produces an average multiplier which is lower than the one estimated by the SLQ and this confirms the conclusion for which the SLQ tends to overestimate impact in comparison to its direct rival. This tendency would depend on the incapability of the SLQ of taking into account the importance of the purchasing sector. Actually, the use of the SLQ implies that a sector with a bigger relative importance than the national one can sell goods and services to a sector, whose relative importance is less than that at a national level, to the extent suggested by the national coefficient, so overestimating the regional input coefficient and thus the value of multipliers (Johns and Leat, 1987).

Graphical results of the application of the MSP over 100 randomly generated national I-O table are shown in Figure 1. The figure provides a geometric representation of the differences existing among the methods.

A first consideration is that explained dispersion is very high, being near 100%. This is a demonstration of goodness of the technique employed and highlights the existence of a common and repeated structure in terms of deviations among the methods.

Table 1: Sector output multipliers estimated through several regionalization methods after 100 iterations

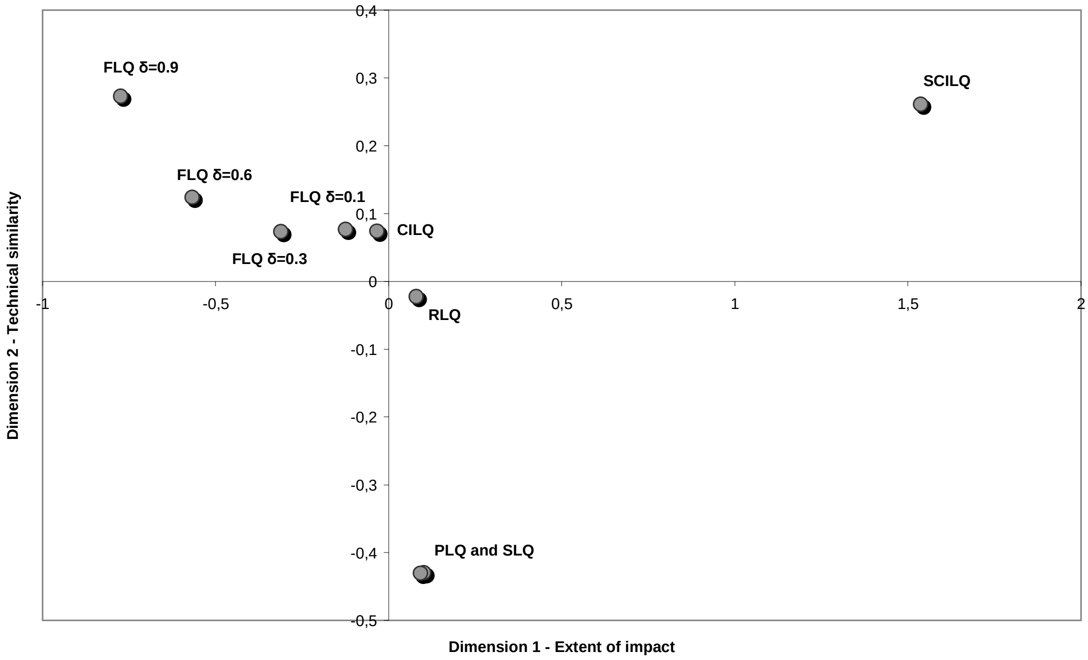
Methods	No hypothesis: Output multiplier (average)	No hypothesis: Rank (increasing sort)	Hypothesis 1*: Output multiplier (average)	Hypothesis 1*: Rank (increasing sort)	Hypothesis 2**: Output multiplier (average)	Hypothesis 2**: Rank (increasing sort)
SLQ	1.622	8	1.646	6	1.718	6
PLQ	1.620	7	1.646	5	1.718	5
CILQ	1.585	5	1.680	8	1.833	8
SCILQ	1.909	9	1.974	9	2.100	9
RLQ	1.611	6	1.682	7	1.802	7
FLQ (δ=0.1)	1.566	4	1.593	4	1.660	4
FLQ (δ=0.3)	1.528	3	1.552	3	1.614	3
FLQ (δ=0.6)	1.472	2	1.491	2	1.542	2
FLQ (δ=0.9)	1.420	1	1.433	1	1.471	1

* intrasectoral coefficients = 20% of total input costs

** intrasectoral coefficients = 40% of total input costs

Source: Author's elaboration

Figure 1: Graphical representation of dissimilarities among regionalization methods, Multidimensional Scaling Procedure (100 iterations, no hypothesis on intrasectoral coefficients)



Measures of Fit – S-Stress: 0.017; Explained dispersion: 0.992

Source: Author's elaboration

With respect to dimension 1, the FLQ, on one hand, and the SCILQ, on the other hand, are located in opposed points. We can also note that, as the value of the parameter δ increases, the FLQ moves on the left. For values of the parameter near zero, the FLQ becomes closer to the CILQ. This can be easily demonstrated mathematically since, when $\delta = 0$, the FLQ coincides with the CILQ. The SLQ, the PLQ, the RLQ and the CILQ are located in a middle position. In particular, the SLQ and the PLQ are in the same position and RLQ is very near the SLQ and the PLQ. Examining the sequence of methods from the left side to the right one and on the basis of the results obtained in terms of multipliers, dimension 1 may be interpreted as extent of impact or tendency to overestimate impact.

Analysing the degree of dissimilarities of methods in respect to dimension 2, two distinct groups can be identified: one regroups the SLQ and the PLQ whereas the other one is composed of the RLQ, the CILQ, the SCILQ and all the versions of the FLQ. Therefore, dimension 2 seems to propound the technical classification of the methods. On one hand, there are methods based on cell-by-cell adjustments (CILQ, RLQ, SCILQ, FLQ) and on the other hand there are methods based on row adjustments (SLQ, PLQ). Definitively, dimension 2 may be referred as an index of technical similarity (in terms of construction) among methods.

The advantage offered by the graphical representation is that it gives an immediate picture about similarity existing among the methods and the direction of impact estimated by different regionalization methods. Trough a look at the graph, policy makers and I-O analysts can know in advance what the choice of a method rather than another brings about in terms of impact estimated. It appears clearly that the SCILQ and the FLQ produce completely different results whereas the use of the SLQ, the PLQ, the CILQ and the RLQ do not cause excessive discrepancy in terms of results.

In order to compare the results with empirical reality, the same application was carried out on a concrete case. As a starting point, we took the 59-sector I-O table constructed in 2000 for Italy (ISTAT, 2006). The national table was then aggregated to 24 sectors corresponding to the common minimum level of sectoral disaggregation allowed by data availability at a regional level. Then, for each of the existing 20 Italian region we applied the 9 regionalization methods illustrated above to obtain 9 corresponding regional I-O matrices. Data on employment used for regionalising the national I-O table come from the ISTAT database (ISTAT, 2005). Totally, 180 regional I-O matrices were calculated. From them, the corresponding Leontief inverses were derived and the relevant sector multipliers were calculated.

Table 2 shows average multipliers associated to the Italian regions, deriving from the application of the regionalization methods. From an economical point of view, it is interesting to note that the majority of the methods are able to capture the structural differences which exist among the Italian regions. This is shown by the relatively high values of the correlation coefficients between multipliers and regional GDP, taken as an indicator of social and economic development.

Table 2: average sector multipliers calculated for regionalization method and Italian region, 2000

Regions	Methods									Range*	VC**	GDP at c.v.
	SLQ	PLQ	CILQ	SCILQ	RLQ	FLQ δ=0.1	FLQ δ=0.3	FLQ δ=0.6	FLQ δ=0.9			
Piemonte	1.638	1.637	1.634	1.749	1.640	1.510	1.314	1.157	1.084	61.3	15.2	100,511
Valle d'Aosta	1.413	1.413	1.474	1.624	1.451	1.253	1.100	1.021	1.004	61.8	16.0	3,082
Lombardia	1.648	1.648	1.620	1.755	1.634	1.551	1.441	1.283	1.166	50.5	11.9	237,200
Trentino-Alto Adige	1.505	1.505	1.565	1.713	1.552	1.361	1.180	1.067	1.023	67.4	16.5	25,254
Veneto	1.603	1.602	1.585	1.667	1.595	1.466	1.290	1.137	1.065	56.5	14.7	106,508
Friuli-Venezia Giulia	1.558	1.557	1.591	1.701	1.584	1.382	1.186	1.064	1.023	66.3	17.0	26,995
Liguria	1.514	1.516	1.600	1.756	1.579	1.392	1.210	1.090	1.046	67.9	16.4	35,250
Emilia Romagna	1.613	1.612	1.592	1.685	1.601	1.467	1.288	1.137	1.066	58.1	14.9	102,801
Toscana	1.584	1.584	1.587	1.671	1.589	1.440	1.255	1.113	1.050	59.1	15.3	79,185
Umbria	1.625	1.625	1.606	1.710	1.614	1.405	1.170	1.049	1.015	68.5	18.2	16,384
Marche	1.550	1.548	1.574	1.649	1.570	1.373	1.174	1.063	1.021	61.5	16.5	29,924
Lazio	1.510	1.513	1.604	1.778	1.580	1.454	1.315	1.179	1.107	60.6	13.9	117,675
Abruzzo	1.560	1.560	1.603	1.670	1.596	1.371	1.157	1.049	1.016	64.4	17.4	21,935
Molise	1.493	1.494	1.556	1.637	1.535	1.290	1.094	1.020	1.005	62.9	17.4	5,105
Campania	1.504	1.506	1.611	1.711	1.581	1.433	1.262	1.119	1.061	61.3	15.0	76,223
Puglia	1.494	1.495	1.579	1.692	1.558	1.402	1.239	1.105	1.049	61.3	15.0	54,808
Basilicata	1.481	1.481	1.578	1.724	1.555	1.346	1.146	1.042	1.012	70.4	17.4	8,664
Calabria	1.408	1.410	1.564	1.730	1.524	1.346	1.206	1.086	1.034	67.3	15.7	25,422
Sicilia	1.488	1.492	1.598	1.761	1.568	1.411	1.266	1.137	1.080	63.1	14.8	67,268
Sardegna	1.502	1.503	1.593	1.770	1.569	1.381	1.216	1.089	1.040	70.2	16.5	24,973
Italy	1.535	1.535	1.586	1.708	1.574	1.402	1.225	1.100	1.048	63.0	15.6	1,165,167
Pearson's coefficient***	0.584	0.590	0.513	0.369	0.598	0.878	0.953	0.971	0.960	-	-	-

* The range is obtained as: [(max - min)/min]*100

** VC = Variation coefficient obtained as a percentage ratio between standard deviation and mean

*** Correlation coefficient calculated between multipliers and regional GDP

Source: Author's elaboration

Among the methods, the FLQ is the one which exhibits the highest correlation (specifically the FLQ with $\delta = 0.6$) and this sanctions its superiority, in comparison with the others, in depicting regional economies. The methods which show the lowest values are the CILQ and, above all, the SCILQ whose Pearson's coefficient is below 0.4.

These methods are the only ones which do not take into account the importance of the regions and this could be the reason why they demonstrate minor economical validity.

With regard to the tendency to over-(under-)estimate impact, the table reveals that most results coming from the simulation are confirmed. Moreover, from the table, policy implications coming from the use of a given regionalisation method appear evident. In effect, the maximum variability in terms of impact deriving from the choice of the methods is particularly high as it is demonstrated by the percentage difference between the minimum impact deriving from the FLQ and the maximum impact estimated by the SCILQ (63%) and by the variation coefficient (15.6%) calculated at a national level. At a regional level, the percentage difference in terms of impact oscillates from a minimum of 50.5% in the case of the Umbria region (located in the Centre Italy) to a maximum of 70.4% in the case of the Basilicata region (located in the South Italy), whereas the variation coefficient goes from 11.9% in the case of the Lombardia region (located in the North Italy) to 18.2% in the case of the Umbria region.

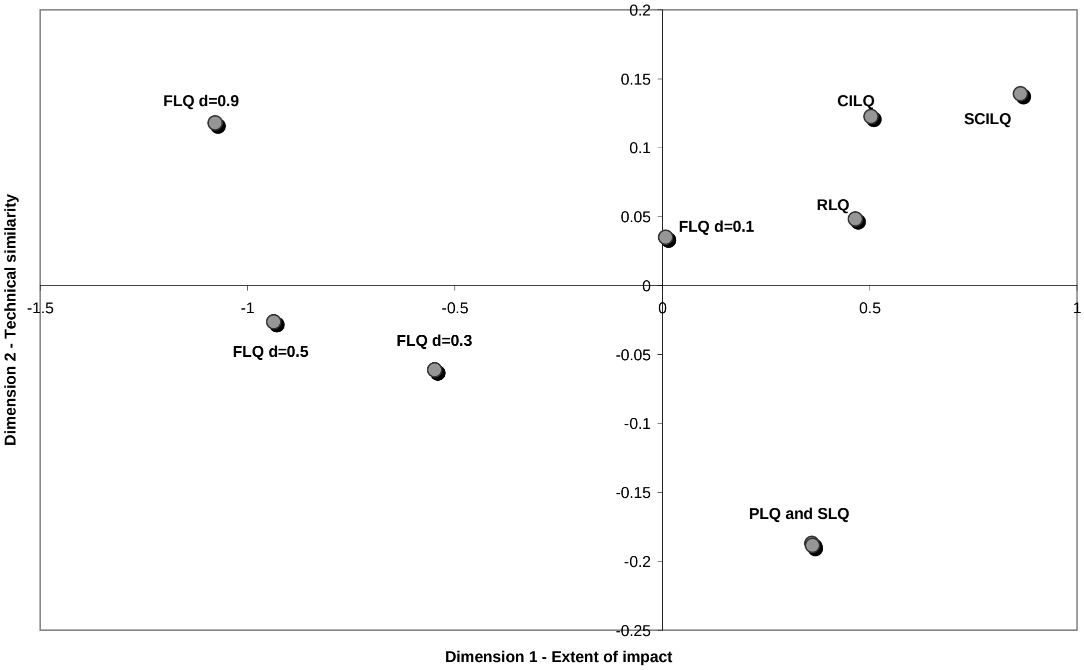
A contrasting result compared to simulation analysis is that the CILQ and the RLQ tend to overestimate impact in comparison with the SLQ and the PLQ. In 15 out of 20 regional cases considered, the CILQ and the RLQ produce higher impact than the SLQ and the PLQ. This is illustrated graphically also in Figure 2 which shows the results coming from the application of the MSP to the distance matrices. Compared to the results of the simulation analysis, we can note that, looking at dimension 1, the RLQ and the CILQ are actually located on the right side of the SLQ and the PLQ.

The reason for this could be searched for in the way the methods treat intrasectoral coefficients. Actually, with reference to the CILQ, it turns out that $CILQ_{ii} \equiv 1$ when instead the SLQ can take inferior values. This brings about that the intraregional input coefficients are overestimated. It is as saying that the selling sector is always able to satisfy all requirements even when the local industry is small (Morrison and Smith, 1974). Moreover, regional intrasectoral coefficients, remaining equal to the national coefficients, wrongly incorporate trade among regions (Flegg et al., 1995). This is the reason why the FLQ's formula imposes the use of the SLQ in correspondence with intrasectoral coefficients. As regards the RLQ, excluding the case in which the SLQ equals one, RLQ_{ii} results to be always bigger than SLQ_i with the consequence that intrasectoral coefficients are overestimated.

This tendency to overestimate intrasectoral coefficients can be reinforced by the fact that input flows along the main diagonal usually represent a conspicuous share of total input costs. In the empirical case considered, on average, coefficients located on the principal diagonal of the national I-O table represent 27% of the respective total input cost. This is also due to the sectoral aggregation, often forced by data unavailability, which is so high that does not allow one to know flows of inputs among branches which are part of the sector.

To investigate this hypothesis, two further Monte Carlo simulations were carried out in which we imposed that the I-O coefficient located on the principal diagonal had to be 20% and 40% of the column sum, respectively.

Figure 2: Graphical representation of dissimilarities among regionalization methods, Multidimensional Scaling Procedure, 20 Italian regions



Measures of Fit – S-Stress: 0.27; Explained dispersion: 0.985

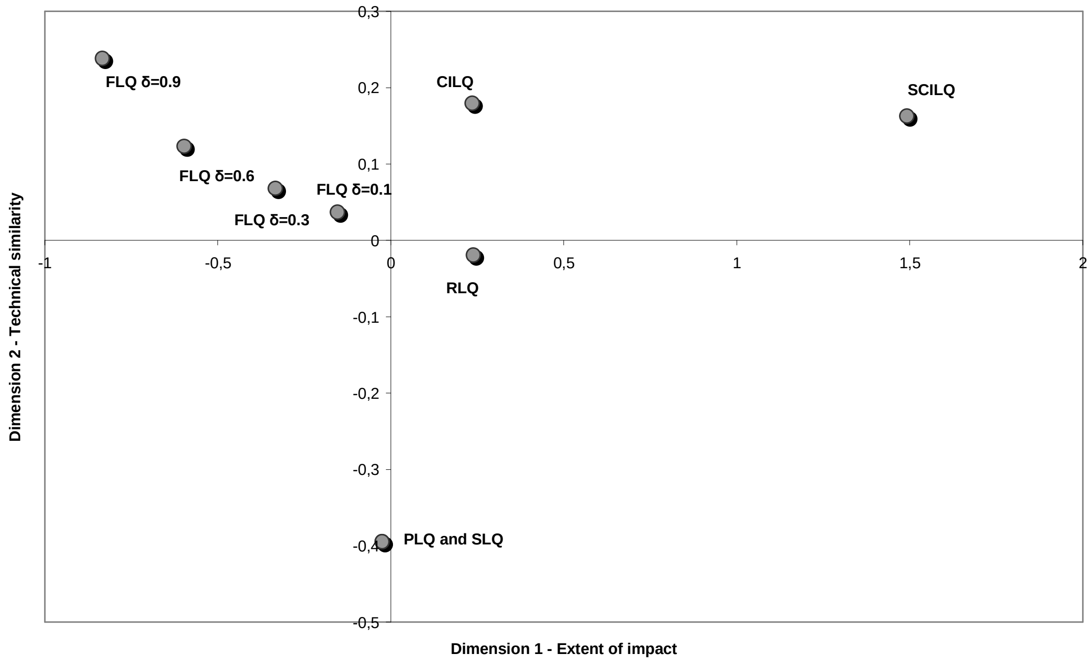
Source: Author's elaboration

Table 1 shows the average value of sector output multipliers and the corresponding ranking under the two hypotheses. The results confirm the tendency of the CILQ and the RLQ to overestimate impact compared to the SLQ. They also show that as the share of intrasectoral coefficients increases, the multiplier derived by the CILQ is bigger than that obtained by the RLQ. This is because, given $SLQ_i < 1$, it results that RLQ_{ii} is always lower than $CILQ_{ii}(=1)$.

Moreover, it appears that multipliers estimated by all the regionalization methods tend to raise as the intrasectoral aggregation increases. This increase is a result of the bias generated by the widest problem of sectoral aggregation (Lahr and Stevens, 2002).

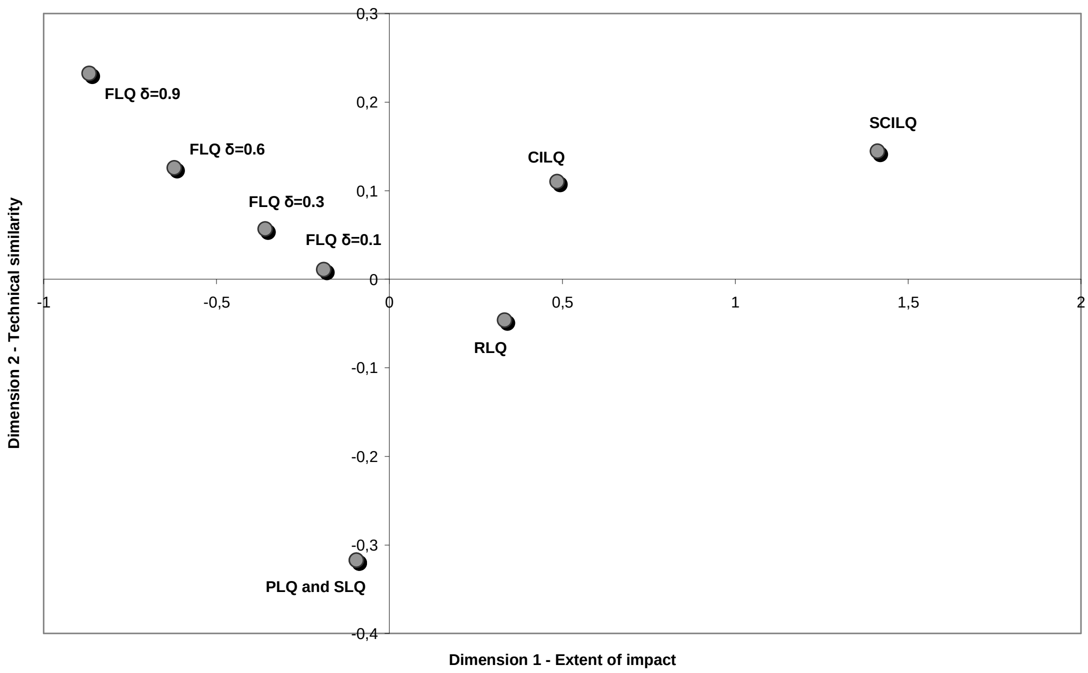
Figures 3 and 4 show the results of the application of the MSP over 100 randomly generated national I-O tables supposing that 20% and 40% of total input costs are absorbed by intrasectoral coefficients, respectively. In both cases, the explained dispersion is very high, demonstrating the goodness of the technique employed and the presence of a common and repeated structure in terms of differences among the methods. From the figures, we can note that the main difference compared to simulation analysis with no hypothesis about intrasectoral coefficients is that the SLQ and the PLQ are located on the left side of the CILQ and the RLQ.

Figure 3: Graphical representation of dissimilarities among regionalization methods, Multidimensional Scaling Procedure (100 iterations, hypothesis 1: intrasectoral coefficients = 20% of total input costs)



Source: Author's elaboration

Figure 4: Graphical representation of dissimilarities among regionalization methods, Multidimensional Scaling Procedure (100 iterations, hypothesis 2: intrasectoral coefficients = 40% of total input costs)



Measures of Fit – S-Stress: 0.020; Explained dispersion: 0.992

Source: Author's elaboration

Moreover, as the level of intrasectoral aggregation increases, the CILQ and the RLQ tend to become farer from the SLQ, confirming the results coming from the analysis of multipliers.

5 Concluding remarks

This paper has investigated the tendency of a battery of widely discussed non-survey methods to estimate impact. For this aim, a Monte Carlo simulation has been carried out to assess this tendency. Successively, a multidimensional scaling procedure has been applied to search for a common and repeated structure of differences among the methods and to give an immediate picture of the features of the methods in terms of extent of impact and technical similarity. To validate the results, the same procedure has been applied to multipliers derived from 2000 I-O tables, which have been mechanically constructed for the 20 Italian regions.

The results have demonstrated that whether a regional I-O table, which is often the basis for deriving SAMs, is constructed by using a non-survey method and this often happens for lack of data at a local level, the choice of the regionalization method becomes crucial. According to the chosen method, estimates of multipliers could be considerably bigger or lower. This can have serious repercussions in terms of policy choices and, therefore, policy makers and analysts should be aware of it. In this respect, to reduce bias, it would be more correct to derive an interval of impact rather than estimating impact punctually.

Moreover, the results have statistically confirmed the existence of a tendency of methods to over-(under)-estimate impact, partly validating the empirical studies found in literature and lighting less investigated methods (i.e. the SCILQ). In particular, it was ascertained that, in comparison to the other methods, the FLQ underestimates impact whereas the SCILQ, for its feature of adjusting national coefficient even upwards, produces very high values. Between these two methods, in a decreasing sort, we found the SLQ, which is considered notoriously as the method which exaggerates the extent of multipliers, followed by the PLQ, the RLQ and the CILQ. However, the results have also shown that sectoral aggregation can reverse this tendency. Specifically, it has emerged that from a certain level of intrasectoral aggregation onward, the CILQ and the RLQ tend to produce higher impact. This has been noticed in a real case represented by the 20 Italian regions in 2000 and has been confirmed statistically. Finally, the correlation analysis between multipliers derived by the methods analysed and regional GDP has shown that the majority of the methods analysed are able to capture overall economical disparities among Italian regions. Among all, the FLQ has demonstrated to be considerably superior whereas those which have shown a minor economical validity are the CILQ and the SCILQ. This last result was attributed to their impossibility of taking into consideration the importance of the regions.

References

Bonfiglio A, Chelli F, 2004, "An Impact Analysis of SAPARD in Rural Areas by Alternative Methods of Regionalization", Quaderni di ricerca 238, Department of Economics, Polytechnic University of the Marche, Ancona, Italy

Brucker S M, Hastings S E, Latham W R, 1987, "Regional input-output analysis: a comparison of five 'ready-made' model systems" *The Review of Regional Studies* **17** 1 - 16

Flegg T A, Webber C D, 1997, "On the Appropriate Use of Location Quotients in Generating Regional Input-Output Tables: Reply" *Regional Studies* **31** 795 - 805

Flegg T A, Webber C D, Elliot M V, 1995, "On the Appropriate Use of Location Quotients in Generating Regional Input-Output Tables: Reply" *Regional Studies* **29** 547 - 561

Greenstreet D, 1989, "A conceptual framework for construction of hybrid regional input-output models" *Socio-Economic Planning Sciences* **23** 283 - 289

Harrigan F J, McGilvray J W, McNicoll I H, 1980, "Simulating the structure of a regional economy" *Environment and Planning A* **12** 927 - 936

ISTAT, 2005 *Conti economici regionali. Anni 1980-2004* [Regional economic accounts. Years 1980-2004] (ISTAT, Rome)

ISTAT, 2006 *Conti nazionali: risorse e impieghi. Il nuovo sistema input output. Anni 1995-2003* [National accounts: supply and use. The new Input-Output system. Years 1995-2003] (ISTAT, Rome)

Jackson R W, 1998, "Regionalizing national commodity-by-industry accounts" *Economic Systems Research* **10** 223 - 238

Jensen R C, 1987, "On the concept of ready-made regional input-output models" *The Review of Regional Studies* **17** 20 - 24

Jensen R C, Mandeville T D, Karunarante N D, 1979 *Regional Economic Planning: Generation of Regional Input-Output Analysis* (Croom Helm, London)

Johns P L, Leat P M K, 1987, "The Application of Modified GRIT Input-Output Procedures to Rural Development Analysis in Grampian Region" *Journal of Agricultural Economics* **38** 245 - 256

Lahr M L, 1993, "A Review of Literature Supporting the Hybrid Approach to Constructing Regional Input-Output Models" *Economic System Research* **5** 277 - 293

Lahr M L, 2001, "A Strategy for Producing Hybrid Regional Input-Output Tables", in *Input-Output Analysis: Frontiers and Extensions* Eds M L Lahr, E Dietzenbacher (Palgrave, London) pp 211 - 242

Lahr M L, Stevens B H, 2002, "A Study of the Role of Regionalization in the Generation of Aggregation Error in Regional Input-Output Models" *Journal of Regional Science* **42**(3) 477 - 507

Lindall S A, Olson D C, 1998 *The IMPLAN Input-Output System* (IMPLAN Group Inc., Stillwater, Minnesota)

Madsen B, Jensen-Butler C, 1999, "Make and Use Approaches to Regional and Interregional Accounts and Model" *Economic Systems Research* **11** 277 - 299

Midmore P, 1991, "Input-output in agriculture: a review", in *Input-Output Models in the Agricultural Sector* Ed P Midmore (Avebury, England, Aldershot) pp 5 - 20

Miller R E, Blair P D, 1985 *Input-Output Analysis: foundations and extensions* (Prentice-Hall, Inc., Englewood Cliffs, New Jersey)

Morrison W I, Smith P, 1974, "Nonsurvey Input-Output Techniques at the Small Area Level: An Evaluation" *Journal of Regional Science* **14** 1 - 14

Oude Wansink M J, Maks J A H, 1998, "Constructing Regional Input-Output Tables", Working Paper, Maastricht University, Maastricht

Round J I, 1972, "Regional Input-Output Models in U.K.: A Reappraisal of Some Techniques" *Regional Studies* **6** 1 - 9

Round J I, 1978, "An Interregional Input-Output Approach to the Evaluation of non-survey methods" *Journal of Regional Science* **18** 179 - 194

Round J I, 1987, "A note on 'Ready-Made' Regional Input-Output Models" *The Review of Regional Studies* **17** 26 - 27

Stevens B H, Treyz G I, Lahr M L, 1989, "On the comparative accuracy of rpc estimating techniques", in *Frontiers of Input-Output Analysis* Eds R E Miller, K R Polenske, A Z Rose (Oxford University Press, New York) pp 245 - 257

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

Treyz G I, Rickman D S, Shao G, 1992, "The REMI Economic Demographic Forecasting and Simulation Model" *International Regional Science Review* **14** 221 - 253

West G R, 1990, "Regional Trade Estimation: A Hybrid Approach" *International Regional Science Review* **13** 103 - 118