UNIVERSITÀ POLITECNICA DELLE MARCHE

DIPARTIMENTO DI ECONOMIA

CAN NON-SURVEY METHODS SUBSTITUTE FOR SURVEY-BASED MODELS? A PERFORMANCE ANALYSIS OF INDIRECT TECHNIQUES OF ESTIMATING I-O COEFFICIENTS AND MULTIPLIERS

ANDREA BONFIGLIO

QUADERNO DI RICERCA n. 230

Marzo 2005

Can Non-survey Methods Substitute for Survey-based Models? A Performance Analysis of Indirect Techniques of Estimating I-O Coefficients and Multipliers

Andrea Bonfiglio*

Abstract

The objective of this article is to evaluate performances of eight non-survey methods in reproducing a survey-based I-O model in a both partitive and holistic sense. In order to evaluate methods, a package of statistics is selected and applied. The main results of the analysis are as follows. The Purchases-only Location Quotient (PLQ) demonstrates to overcome all the others in reproducing survey-based I-O coefficients whereas the Flegg *et al.* Location Quotient (FLQ) performs better in estimating survey-based output multipliers. Overall, the non-survey methods examined produce better results in estimating multipliers rather than I-O coefficients. In any case, estimates are too far from the survey-based ones. For this reason, methods should not be used alone but integrated with all available exogenous information within hybrid procedures.

Keywords: survey-based model, non-survey techniques, input-output matrices, performances analysis

J.E.L. Classification: C15, C42, C67, R15

* Department of Economics, Faculty of Economics "G. Fuà", Polytechnic University of the Marche, Piazzale Martelli, 8, 60121 Ancona. Tel.: +39 071 2207116, Fax.: +39 071 2207102, e-mail: bonfiglio@dea.unian.it. I am grateful to an anonymous referee for remarks and suggestions on a previous version of the paper.

Contents

1 Introduction

Construction of regional input-output tables still represents an important task for regional analysts. The main advantage offered by regional I-O models is the possibility of measuring impact in the regional economy from local or national policy at a high level of sector disaggregation. But this is not the only advantage related to construction of regional I-O models. A further advantage is that the entire modelling process, including data development, may improve the knowledge of a regional economy. Moreover, regional I-O tables represent the basis for other model applications and extensions. A clear example is given by social accounting matrices (SAMs) which are natural extensions of regional input-output tables and a precious databank for computable general equilibrium models (Gerking *et al.*, 2001).

Unfortunately, construction of an input-output table implies 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. That requires the collection of a great volume of information, which, at a sub-national level, is difficulty 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: “survey”, “non-survey” and “hybrid” approaches.

While the 1960s have been dominated by survey-based models (Richardson, 1985), in the next years, efforts to produce this kind of models have considerably diminished. 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), fundamentally based on non-survey techniques, are the most used by I-O analysts since they permit to reduce considerably costs associated to survey-based models and, especially in the case of hybrid techniques, to reach satisfactory levels of reliability.

A still wide 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 real performances in representing in detail a given productive structure (partitive accuracy) or in estimating impacts as faithfully as possible (holistic accuracy) (Jensen, 1980). In spite of some possible drawbacks[2], an usually used method by analysts to evaluate performances of indirect methods is to measure deviations between indirectly constructed tables and survey-based tables.

In this connection, there is sufficient agreement in literature on stating that hybrid methods perform better than non-survey techniques. Several empirical studies have in fact demonstrated the superiority of hybrid techniques (like RAS) in generating closer coefficients and multipliers to the survey-based ones (Morrison and Smith, 1974; Sawyer and Miller, 1983; Willis, 1987; Strassoldo, 1988; Harris and Liu, 1998). Other studies have proved that as the quantity of exogenous information inserted in the table raises, improvements tend to increase (Dewhurst, 1992; Lahr, 2001).

Instead, with reference to non-survey methods, there would seem there is no peaceful agreement on the method which can be defined as better than others. Different methods have been labelled as more effective on the basis of empirical results: the Simple Location Quotient (Schaffer and Chu, 1969a, 1969b; Morrison and Smith, 1974; Eskelinen and Suorsa, 1980; Sawyer and Miller, 1983); the Supply Demand Pool technique

[2] All comparisons between survey-based models and indirectly constructed matrices start from the assumption that the former generate true values. This assumption may be wrong owing to the possibility of sampling errors as well as reconciliation adjustments, indirect estimation of some quantities and various arbitrary procedures used in constructing survey-based tables (Richardson, 1985).

(Strassoldo, 1988)[3]; the Flegg *et al.* Location Quotient (Flegg and Webber, 2000).

These discordant conclusions, which represent one of the interest reasons in analysing non-survey methods, may depend on the fact that existing empirical studies compare different batteries of techniques and use diverse types of statistics to measure the distance among indirect and survey-based estimates.

However, interest in non-survey techniques mostly comes from three other grounds. Firstly, in spite of several criticisms addressed to mechanical procedures (Round, 1987), non-survey techniques have been redeemed thanks to the diffusion and a wide use of ready-made models, which are fundamentally based on this type of techniques. Therefore, many models of impact prediction and evaluation are definitively based on non-survey techniques. Secondly, starting from the middle of the nineties, new versions of non-survey methods have appeared, attesting a ceaseless interest in shortcuts to derive regional I-O tables (Flegg *et al.* 1995; Flegg and Webber, 1997; Oude Wansink and Maks, 1998). Thirdly, non-survey techniques are widely used 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 is likely to be affected by the kind of non-survey method used to derive preliminary estimates. More the non-survey method is reliable more the final I-O table will be closer to a "real" I-O table.

[3] Strassoldo reports and comments on empirical results obtained in a Mogorovich's (1987) study.

The objective of this research is to attempt to evaluate performances of eight non-survey methods in reproducing a survey-based I-O model in a both partitive and holistic sense. This work is articulated as follows. The second section is finalised to describe the selection process of statistics used to carry out a performance analysis. The next section shows results of an empirical analysis. Finally, the last section provides some concluding remarks.

2 Choosing a battery of statistics for comparison

In the literature, many measures have been used to estimate deviations between survey-based and indirect-method-based tables (Lahr, 2001). Since measures possess different properties and characteristics, one of the problems associated to empirical studies on performances of indirect techniques is that results could be affected by the statistic chosen as a mean of comparison. To reduce the risk that the choice of the statistic used can affect conclusions hardly, several statistics should be jointly employed (Butterfield and Mules, 1980). However, some of them could share common characteristics, measuring a similar pattern of simulation error. Consequently, the joint use of these statistics could alter conclusions by making a given method better or worse than it really is.

To avoid these drawbacks, the solution adopted was to identify a package of statistics which had to be both as wide as possible to diminish the risk that conclusions may be influenced by the statistic used and net of redundant measures which would have had the effect of distorting results on performances of methods. Towards this aim, 35 statistics (of which some are just variants of same statistics) found in the literature were first explored. They were compared to each other by a correlation analysis and those having similar characteristics were excluded from the analysis of performances of indirect techniques.

2.1 A brief review of statistics for comparing survey-based with indirectly constructed I-O matrices

The measures used to compare survey-based with indirectly constructed I-O tables can be broadly classified intro three main categories: traditional statistics, general distance statistics and information-based statistics.

Within the class of traditional statistics, there can be included the following ones: the chi-square statistic (CS); the weighted chi-square statistic (WCS); regression and correlation analysis measures (intercept, slope and squared correlation coefficient, R^2); the Euclidean metric distance (EMD).

The general distance statistics are more numerous and they are: the index of inequality (U) and the related bias (UM), variance (US) and covariance (UC); the index of relative change (RC); the similarity index (SI); the absolute mean difference (AMD); the mean absolute difference as a percentage of the mean coefficient (MPMC); the absolute standard deviation difference (ASDD); the mean absolute difference (MAD); the mean absolute relative difference (MARD); the mean weighted absolute error (MWAE); the mean weighted error (MWE); the MWE along with the related variance (MWEV); the mean weighted relative error (MWRE); the MWRE along with the related variance (MWREV); the standardized total percentage error (STPE); the weighted absolute difference (WAD); the coefficient of equality (CE) used through its related mean (MCE), standard deviation (SDCE), maximum value (MXCE) and degree of approximation (DA).

Finally, within the class of information-based statistics, there can be included the information content index (I).[4]

[4] It is recognized that some measures could have been neglected. Therefore, we apologize with all those authors and researchers whose measures analysed do not appear in the list.

The measures abovementioned have been widely discussed in the literature. For this reason, only a brief description will be here given.

The CS (Schaffer and Chu, 1969a; Morrison and Smith, 1974; Strassoldo, 1988) takes the following form:

$$\chi^2 = \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\left(a_{ij} - b_{ij}\right)^2}{b_{ij}} \tag{1}$$

where a_{ij} are estimated coefficients whereas b_{ij} are survey-based coefficients. Obviously, the lower the chi-square the better the estimate is. The main problem, which characterises most statistics based on relative ratios, is that when b_{ij} is null, the standardization is meaningless.

The WCS is a modified version of CS (Flegg and Webber, 2000). It takes the form:

$$WCS = \sum_{j=1}^{n} w_j \sum_{i=1}^{n}\frac{\left(a_{ij} - b_{ij}\right)^2}{b_{ij}} \tag{2}$$

It is weighted on the basis of proportions of employment of purchasing sector j (w_j) in order to take account of the relative importance of sectors.

The regression and correlation approach (Schaffer and Chu, 1969b; Butterfield and Mules, 1980; Morrison and Smith, 1974; Strassoldo, 1988; Harris and Liu, 1998; Gilchrist and Louis, 1999) is based on the well-known regression equation:

$$a_{ij} = \alpha + \beta b_{ij} \tag{3}$$

From this equation, three test coefficients can be calculated: (i) the squared correlation coefficient (R^2) which should be close to unity for a good fit; (ii) an intercept or constant term (α) which should be close to zero for a good fit and (iii) the slope coefficient or regression coefficient (β) which should be close to unity for a good fit. Calculation of tests can be made on columns (R^2-C, intercept-C, slope-C), rows (R^2-R, intercept-R, slope-R) or on the whole matrix (R^2-A, intercept-A, slope-A).

The EMD (Harrigan *et al.,* 1980; Robinson *et al.*, 2001) also known as root mean square takes the following form:

$$EMD = \frac{1}{n^2}\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}\left(a_{ij} - b_{ij}\right)^2} \tag{4}$$

The Theil's U (Theil, 1971; Stevens and Trainer, 1976; Jackson and Murray, 2004), commonly called Theil's index of inequality, takes the following form:

$$U = \sqrt{\frac{\sum_{i=1}^{n}\sum_{j=1}^{n}\left(a_{ij} - b_{ij}\right)^2}{\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}^2}} \tag{5}$$

This measure provides an estimate of overall distance proportion. It ranges from 0 to 1. When U=0, a perfect fit is obtained. From this measure, further three proportions can be calculated: bias, U^M, variance, U^S, and covariance U^C.

$$U^M = \frac{\left(\bar{a} - \bar{b}\right)^2}{\left(1/n^2\right)\sum_{i=1}^{n}\sum_{j=1}^{n}\left(a_{ij} - b_{ij}\right)^2},\; U^S = \frac{\left(\sigma_a - \sigma_b\right)^2}{\left(1/n^2\right)\sum_{i=1}^{n}\sum_{j=1}^{n}\left(a_{ij} - b_{ij}\right)^2},$$

$$U^C = \frac{2\left(1 - r\right)\sigma_a\sigma_b}{\left(1/n^2\right)\sum_{i=1}^{n}\sum_{j=1}^{n}\left(a_{ij} - b_{ij}\right)^2}. \tag{6}$$

where $\overline{a}$ is the mean of a_{ij}; $\overline{b}$ is the mean of b_{ij}; σ_a is the standard deviation of a_{ij}; σ_b is the standard deviation of b_{ij}; and $r = \dfrac{\left(1/n^2\right)\sum_{i=1}^{n}\sum_{j=1}^{n}\left(a_{ij} - \overline{a}\right)\left(b_{ij} - \overline{b}\right)}{\sigma_a \sigma_b}$.

The bias proportion measures the extent to which the average values of the simulated and actual values deviate from each other. It thus provides an indication of systematic error. Ideally, the bias proportion should be zero. The variance proportion indicates the ability of the model to replicate the degree of variability of the survey-based coefficients. Also in this case, a close variance to zero indicates better estimates. The covariance proportion measures unsystematic error. As simulation performances improve, this proportion approaches one. The optimal distribution of the inequality proportions should be then for any $U > 0$: $U^M = U^S = 0$ and $U^C = 1$.

The RC (Schaffer and Chu, 1969b) takes the following form:

$$RC_{ij} = \frac{\left|a_{ij} - b_{ij}\right|}{\frac{1}{2}\left(a_{ij} + b_{ij}\right)} \tag{7}$$

It ranges from 0 to 2. As the RC approaches 0, simulation improves.

The SI (Morrison and Smith, 1974) is a modified version of RC ranging from zero to unity. It takes the form:

$$SI_{ij} = 1 - \frac{\left|a_{ij} - b_{ij}\right|}{\left(a_{ij} + b_{ij}\right)} \tag{8}$$

Closer values to the unity identify better estimates.

The AMD is an attempt of formal conceptualization of an analysis developed by Jalili (2000)[5]. It takes the following form:

$$AMD = \left|\overline{a} - \overline{b}\right| \tag{9}$$

where $\overline{a}$ and $\overline{b}$ have been previously defined.

The MPMC (Sawyer and Miller, 1983) measures the percentage deviation of the mean coefficient and takes the following form:

$$MPMC = 100 \cdot \left|\frac{\overline{a} - \overline{b}}{\overline{b}}\right| \tag{10}$$

The ASDD is a further attempt of formalization of Jalili's (2000) argumentations[6]. It takes the form:

$$ASDD = \left|\sigma_a - \sigma_b\right| \tag{11}$$

where σ_a and σ_b have been previously defined.

The MAD (Morrison and Smith, 1974; Butterfield and Mules, 1980; Sawyer and Miller, 1983; Strassoldo, 1988; Jalili, 2000; Okuyama *et al.*, 2000) has the following form:

$$MAD = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left|a_{ij} - b_{ij}\right| \tag{12}$$

---

[5] Jalili calculates the means of estimated coefficients using indirect methods and the mean of survey-based coefficients. Then, comparing the means, he defines the better methods as those methods having closer means to the survey-based one. This approach can be formalized calculating the absolute difference between the mean of estimated coefficients and the mean of survey-based coefficients (AMD) and sorting methods on the basis of the extent of AMD. The closer AMD to zero, the better the estimate is.

[6] See note 5 for an explanation, replacing means with standard deviations.

The MARD (Sawyer and Miller, 1983) is a variant of MAD that considers the relative dimension of coefficients. It takes the following form:

$$MARD = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left|\frac{a_{ij} - b_{ij}}{b_{ij}}\right| \tag{13}$$

The MWAE (Flegg and Webber, 2000) takes this form:

$$MWAE = \frac{1}{n}\sum_{j=1}^{n} w_j \sum_{i=1}^{n}\left|a_{ij} - b_{ij}\right| \tag{14}$$

It is very similar to MAD with the difference that employment ratios are used as weights and it is less sensitive to the size of the input-output matrices under study.

The MWE (Flegg and Webber, 2000) takes the following form:

$$MWE = \frac{1}{n}\sum_{j=1}^{n} w_j \sum_{i=1}^{n}\left(a_{ij} - b_{ij}\right) \tag{15}$$

One problem related to this statistic is that negative and positive values can compensate each other giving a misleading impression on goodness of a method.

The MWEV has been introduced to take account of both MWE's properties and the intersectoral variation of simulation errors (Flegg and Webber, 2000). The statistic has the following form:

$$\omega_{MWE} = (MWE - 0)^2 + \sigma^2_{MWE} \tag{16}$$

where σ^2_{MWE} is the variance of the values $w_j \sum_{i=1}^{n}\left(a_{ij} - b_{ij}\right)$ with $j = 1,\dots,n$. By minimizing MWEV rather than MWE, both bias and variance are considered.

The MWRE (Flegg and Webber, 2000) takes the following form:

$$MWRE = \frac{1}{n}\sum_{j=1}^{n} w_j \frac{\sum_{i=1}^{n}\left(a_{ij} - b_{ij}\right)}{\sum_{i=1}^{n} b_{ij}} \tag{17}$$

It was conceived to overcome limits of MWE taking account of two factors: (a) the relative size of the simulation error for each coefficient, calculating $\left(a_{ij} - b_{ij}\right)/b_{ij}$; (b) the relative size of the coefficient in question, calculating $b_{ij} \Big/ \sum_{i=1}^{n} b_{ij}$.

The MWREV (Flegg and Webber, 2000) incorporates the properties of MWRE and takes account of intersectoral variation in simulation error. It is:

$$\omega_{MWRE} = (MWRE - 0)^2 + \sigma^2_{MWRE} \tag{18}$$

where σ^2_{MWRE} is the variance of the values $w_j \sum_{i=1}^{n}\left(a_{ij} - b_{ij}\right) \Big/ \sum_{i=1}^{n} b_{ij}$, with $j = 1,\dots,n$.

The STPE (Miller and Blair, 1983; Jalili, 2000; Canning and Wang, 2004; Jackson and Murray, 2004) has the following structure:

$$STPE = 100 \cdot \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}\left|a_{ij} - b_{ij}\right|}{\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}} \tag{19}$$

It measures the overall percent relative distance in absolute terms between indirect-method-based matrix and the survey-based one.

The WAD (Lahr, 2001; Jackson and Murray, 2004) takes the following form:

$$WAD = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}\left|a_{ij} - b_{ij}\right|\left(a_{ij} + b_{ij}\right)}{\sum_{i=1}^{n}\sum_{j=1}^{n}\left(a_{ij} + b_{ij}\right)} \tag{20}$$

It has been conceived to solve two common problems associated to other measures: low sensitivity to errors in large cells and meaningless standardization. The first problem is faced by adding the term $\left(a_{ij} + b_{ij}\right)$ which weights the absolute difference term in such a way that the errors of large cells are emphasized. As for the second problem, the statistic is never undefined if either of the matrices is null for a cell.

The coefficient of equality (Jalili, 2000) assumes the following form:

$$CE_{ij} = \frac{a_{ij}}{b_{ij}} \tag{21}$$

From this coefficient, the related mean (MCE), standard deviation (SDCE) and maximum value (MXCE) are calculated as follows:

$$MCE = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} CE_{ij}\,;\; SDCE = \sqrt{\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(CE_{ij} - MCE\right)^2}\,;$$

$$MXCE = \max_{i,j} CE_{ij}\,; \tag{22}$$

The closer the MCE and the MXCE to one value, the better the estimate is. The closer the SDCE to zero, the better the simulation is.

The degree of approximation (Jalili, 2000) is obtained as:

$$DA_{ij} = 1 - CE_{ij} \tag{23}$$

It was used to identify the number of coefficients estimated falling into a given percent interval of simulation error.

The information content index (Czamanski and Malizia, 1969; Schaffer and Chu, 1969b; Morrison and Smith, 1974; Strassoldo, 1988) was borrowed from the field of information theory. The survey-based matrix is considered as a forecast of the non-survey estimate and the additional information contained in the latter is measured by the following index:

$$I(A:B) = \sum_{i=1}^{n}\sum_{j=1}^{n}\left| a_{ij} \log_2 \frac{a_{ij}}{b_{ij}} \right| \tag{24}$$

The lower the information content thus calculated, the closer the estimate.

2.2 Examining similarity among statistics through a correlation analysis

The statistics described in the previous section were compared to each other by a correlation analysis in order to identify a representative group of statistics. The main criteria used to select statistics were the following ones: (a) no correlation to others; (b) high and positive correlation (≥ 0.8) to a bigger number of measures; (c) for groups of statistics correlated to each other, higher correlation to the others; (d) for pairs of correlated statistics, analyst's discretion.

The analysis was applied to a set of data whose variables were the statistics examined and the observations were results of application of statistics[7] to random benchmark 100-sector matrices compared to random matrices simulating indirectly constructed 100-sector matrices. 1,000 benchmark matrices were randomly generated[8]. Every benchmark matrix was compared to 1,000 randomly-generated matrices[9] to derive 1,000 corresponding observations for every statistic. Definitively, there were estimated as many correlation matrices as the benchmark matrices were. To synthesise information, there was calculated an average correlation matrix.[10]

Results from correlation analysis are shown in Tab. 1. As can be noted, R^2-R is correlated to no other. Therefore, this statistic was included in the group of statistics to be used.

MAD, MWAE and STPE are highly and positively correlated to a bigger number of measures than the other statistics. Besides being correlated to each other, they are correlated to intercept-C, intercept-A, EMD, U, WAD, INFC. However, MAD and STPE exhibit an identical and higher average correlation to the others than MWAE. Therefore the

[7] Since the scale of statistics examined is different, to avoid reach misleading conclusions, all statistics were early adjusted to the same scale so as to identify better estimates, as the value approaches zero. Adjustments were as follows: as for slope-R, slope-C, slope-A, US, R^2-R, R^2-C, R^2-A, MCE, MXCE, there was calculated an absolute value of one less the value of the statistics; as for SI, there was calculated one less the value of the statistic; as for intercept-R, intercept-C, intercept-A, MWE, MWEV, MWRE, MWREV and DA, there was calculated the absolute value of the statistics.

[8] Some statistics (WCS, MWAE, MWE, MWEV, MWRE, MWREV) require the knowledge of employment weights (w_j). 1,000 of these latter were randomly generated per each benchmark matrix.

[9] Considering that some statistics are not defined for zero cells, to avoid obtaining statistically insignificant results, it was imposed that randomly generated coefficients could not take null values. Moreover, in order to simulate the structure of an I-O matrix, it was imposed that the sum of column coefficients had to be less than unity.

[10] The same experiment was repeated more times obtaining very similar results.

choice could fall on either MAD or STPE. Arbitrarily, STPE, which has the advantage to express proportionate error, was chosen and 10 statistics were excluded from the analysis.

UM, UC, AMD, MPMC, MWE, MWEV, MWRE, MWREV result to be highly correlated to each other. Of these, MWE presents an higher average correlation to the others. Therefore, MWE was selected and this allowed eliminating further 7 measures.

CS and SDCE result to be correlated to both each other and WCS, MARD, MCE and MXCE and DA. However, presenting an higher average correlation to the other measures in comparison with SDCE, CS was selected and its selection permitted to eliminate further 6 statistics.

Finally, couples of highly correlated measures are: slope-C and slope-A, R^2-C and R^2-A, slope-R and intercept-R, US and ASDD, RC and SI. The chosen statistics were slope-C, R^2-C, slope-R, US and SI.

Finally, on 35 initial statistics, 9 statistics were selected. Four traditional statistics (R^2-R, slope-R, R^2-C, slope-C) and five general distance statistics (CS, MWE, STPE, US, SI).

Tab. 1 – Average correlation matrix of the statistics used to compare indirect-technique-based with survey-based I-O tables

	Statistics	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	CS	1.00																
2	WCS	0.94	1.00															
3	SLOPE-C	0.23	0.23	1.00														
4	INTERCEPT-C	0.31	0.30	0.60	1.00													
5	R^2-C	0.00	0.00	0.08	0.06	1.00												
6	SLOPE-R	0.13	0.13	0.03	0.02	0.00	1.00											
7	INTERCEPT-R	0.23	0.23	0.03	0.41	0.01	0.82	1.00										
8	R^2-R	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00									
9	SLOPE-A	0.23	0.23	1.00	0.60	0.08	0.03	0.03	0.00	1.00								
10	INTERCEPT-A	0.31	0.30	0.60	1.00	0.06	0.02	0.41	0.00	0.60	1.00							
11	R^2-A	0.00	0.01	0.00	0.03	0.90	0.01	0.02	0.00	0.00	0.03	1.00						
12	EMD	0.30	0.29	0.77	0.81	0.08	0.03	0.24	0.00	0.77	0.81	0.02	1.00					
13	U	0.30	0.29	0.77	0.81	0.08	0.03	0.24	0.00	0.77	0.81	0.02	1.00	1.00				
14	UM	0.00	0.00	-0.04	-0.03	-0.01	0.00	0.00	0.00	-0.04	-0.03	0.00	0.01	0.01	1.00			
15	US	-0.01	-0.01	-0.04	-0.07	0.00	0.00	0.00	0.00	-0.04	-0.07	0.00	-0.04	-0.04	0.46	1.00		
16	UC	0.00	-0.01	-0.04	-0.04	-0.01	0.00	0.00	0.00	-0.04	-0.04	0.00	0.00	0.00	0.97	0.66	1.00	
17	RC (mean)	0.08	0.08	0.62	0.09	0.04	0.01	-0.17	0.00	0.62	0.08	-0.02	0.54	0.54	0.03	0.01	0.02	1.00
18	SI (mean)	0.08	0.08	0.62	0.09	0.04	0.01	-0.17	0.00	0.62	0.08	-0.02	0.54	0.54	0.03	0.01	0.02	1.00
19	AMD	0.01	0.01	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.05	0.05	0.94	0.42	0.91	0.05
20	MPMC	0.01	0.01	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.05	0.05	0.94	0.42	0.91	0.05
21	ASDD	0.00	0.00	-0.01	-0.04	0.00	0.00	0.02	0.00	-0.01	-0.04	0.00	0.01	0.01	0.43	0.94	0.61	0.03
22	MAD	0.30	0.29	0.77	0.86	0.06	0.02	0.27	0.00	0.77	0.86	0.02	0.97	0.97	0.01	-0.04	-0.01	0.55
23	MARD	0.95	0.90	0.25	0.34	0.00	0.13	0.24	0.00	0.25	0.34	0.01	0.30	0.30	0.00	-0.01	0.00	0.08
24	MWAE	0.29	0.29	0.76	0.85	0.06	0.02	0.27	0.00	0.76	0.85	0.02	0.97	0.97	0.01	-0.04	-0.01	0.54
25	MWE	0.02	0.01	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.06	0.06	0.94	0.42	0.90	0.05
26	MWEV	0.02	0.02	0.00	0.04	0.00	0.00	0.03	0.00	0.00	0.04	0.00	0.08	0.08	0.96	0.43	0.92	0.04
27	MWRE	0.03	0.02	0.00	0.05	0.00	0.00	0.04	0.00	0.00	0.05	0.00	0.08	0.08	0.93	0.41	0.90	0.03
28	MWREV	0.05	0.04	0.00	0.11	0.00	0.00	0.08	0.00	0.00	0.12	0.01	0.12	0.12	0.94	0.41	0.90	0.01
29	STPE	0.30	0.29	0.77	0.86	0.06	0.02	0.27	0.00	0.77	0.86	0.02	0.97	0.97	0.01	-0.04	-0.01	0.55
30	WAD	0.21	0.21	0.76	0.58	0.09	0.02	0.09	0.00	0.77	0.58	0.00	0.93	0.93	0.03	-0.01	0.02	0.68
31	MCE	0.95	0.89	0.25	0.34	0.00	0.13	0.25	0.00	0.25	0.34	0.01	0.30	0.30	0.00	-0.01	-0.01	0.07
32	SDCE	0.89	0.82	0.12	0.15	0.00	0.09	0.13	0.00	0.12	0.15	0.01	0.14	0.14	0.00	0.00	0.00	0.04
33	MXCE	0.81	0.75	0.09	0.11	0.00	0.07	0.10	0.00	0.09	0.11	0.01	0.11	0.11	0.00	0.00	0.00	0.03
34	DA (mean)	0.95	0.89	0.25	0.34	0.00	0.13	0.25	0.00	0.25	0.34	0.01	0.30	0.30	0.00	-0.01	-0.01	0.07
35	I	0.41	0.40	0.46	0.94	0.03	0.05	0.46	0.00	0.46	0.94	0.03	0.77	0.77	-0.02	-0.06	-0.03	0.07

Note: the matrix was calculated on 1,000 correlation matrices obtained comparing 1,000 randomly generated 100-sector benchmark matrices with 1,000,000 randomly generated 100-sector matrices simulating indirectly constructed matrices (1,000 per each benchmark matrix)

Source: Author's elaboration

Tab. 1 – Average correlation matrix of the statistics used to compare indirect-technique-based with survey-based I-O tables (continued)

	Statistics	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
1	CS																		
2	WCS																		
3	SLOPE-C																		
4	INTERCEPT-C																		
5	R^2-C																		
6	SLOPE-R																		
7	INTERCEPT-R																		
8	R^2-R																		
9	SLOPE-A																		
10	INTERCEPT-A																		
11	R^2-A																		
12	EMD																		
13	U																		
14	UM																		
15	US																		
16	UC																		
17	RC (mean)																		
18	SI (mean)	1.00																	
19	AMD	0.05	1.00																
20	MPMC	0.05	1.00	1.00															
21	ASDD	0.03	0.41	0.41	1.00														
22	MAD	0.55	0.05	0.05	0.00	1.00													
23	MARD	0.08	0.01	0.01	0.00	0.32	1.00												
24	MWAE	0.54	0.05	0.05	0.00	0.99	0.31	1.00											
25	MWE	0.05	0.99	0.99	0.41	0.05	0.01	0.05	1.00										
26	MWEV	0.04	0.91	0.91	0.40	0.08	0.02	0.08	0.91	1.00									
27	MWRE	0.03	0.99	0.99	0.40	0.08	0.03	0.08	1.00	0.92	1.00								
28	MWREV	0.01	0.89	0.89	0.39	0.12	0.05	0.13	0.90	0.99	0.91	1.00							
29	STPE	0.55	0.05	0.05	0.00	1.00	0.32	0.99	0.05	0.08	0.08	0.12	1.00						
30	WAD	0.68	0.07	0.07	0.03	0.88	0.20	0.87	0.07	0.08	0.07	0.10	0.88	1.00					
31	MCE	0.07	0.01	0.01	0.00	0.32	1.00	0.31	0.01	0.02	0.03	0.05	0.32	0.20	1.00				
32	SDCE	0.04	0.01	0.01	0.00	0.14	0.88	0.14	0.01	0.01	0.01	0.02	0.14	0.10	0.88	1.00			
33	MXCE	0.03	0.01	0.01	0.00	0.11	0.79	0.11	0.01	0.01	0.01	0.02	0.11	0.08	0.79	0.97	1.00		
34	DA (mean)	0.07	0.01	0.01	0.00	0.32	1.00	0.31	0.01	0.02	0.03	0.05	0.32	0.20	1.00	0.88	0.79	1.00	
35	I	0.07	0.01	0.01	-0.03	0.82	0.44	0.82	0.02	0.05	0.06	0.13	0.82	0.54	0.45	0.21	0.16	0.45	1.00

Note: the matrix was calculated on 1,000 correlation matrices obtained comparing 1,000 randomly generated 100-sector benchmark matrices with 1,000,000 randomly generated 100-sector matrices simulating indirectly constructed matrices (1,000 per each benchmark matrix)

Source: Author's elaboration

3 Empirical evidence

For empirical purposes, the benchmark used was the 44-sector 1974 survey-based I-O table constructed for the Italian Marche region (Santeusanio, 1979). The survey-based I-O matrix was compared to matrices obtained from the 1974 44-sector national I-O table[11] (ISTAT, 1977) through the application of the following non-survey methods: the Simple Location Quotient (SLQ), the Purchases-only Location Quotient (PLQ), the West Location Quotient (WLQ); the Cross-Industry Location Quotient (CILQ), the Semilogarithmic Location Quotient (RLQ), the Symmetric Cross Industry Location Quotient (SCILQ), the Flegg *et al.* Location Quotient (FLQ) and the Supply Demand Pool technique (SDP). For an illustration of these methods see West (1980), Miller and Blair (1985), Flegg *et al.* (1995), Flegg and Webber (1997), Oude Wansink and Maks (1998).[12]

---

[11] Both regional survey-based and national I-O tables, constructed on the basis of ESA1970, are valued at current and purchasers' prices. The former reports domestic flows exchanged within the region. The latter reports total flows, both imported and produced domestically.

[12] To be applied, all the methods considered require sector output or value added data at a regional level. Though it would have been possible to exploit information from the survey-based I-O table, we supposed not to have these data, this being a prevalent situation for I-O analysts, and employment data were used as a proxy. 1974 employment data were taken from the National Institute of Statistic's 1977 publication on employees by economic activity and region from 1970 to 1976 (ISTAT, 1977). Data were not available for the all the 44 industries represented into the I-O tables but only for some industries and groups of industries. Data related to these latter were disaggregated using employment weights deriving from 1971 Census of Industry and Services. With regard to the methods analysed, some further specifications should be made. As for WLQ, we supposed to have data on regional sector consumption, collecting them from the survey-based I-O table. Data on population come from 1971 General Census of population. As for SDP, local final demand (i.e. total final demand net of exports) at a regional level was estimated applying employment ratios to national net final demand. All methods were applied to the national technology matrix (taking account of import and domestic flows) to obtain estimates of regional input coefficients (referring to regional domestic production).

Results of the performances analysis are shown in Tabs. 2, 3, 4, 5 and Figs. 1 and 2.

With regard to ability of methods to replicate survey-based I-O coefficients, it emerges that values of statistics are mostly far from the optimal ones (Tab. 2). Therefore, all methods would not seem to produce very satisfactory results. For instance, slope's values are generally more than 2 times their ideal values; the squared correlation coefficients do not show an evident linear relationship between estimated and survey-based coefficients; SI never overcomes 0.7 for all methods; moreover, STPE points out that all methods produce coefficients which are overall far from the "true" coefficients, in absolute terms, by over 100% reaching the percentage of almost 200% in the case of SCILQ.

Sorting methods on the basis of the values assumed by the 9 selected statistics, it emerges that PLQ obtains the highest average rank (1.4) whereas SCILQ ranks in the last position with an average rank of 7.3 (Tab. 3).

Tab. 2 – Performances of non-survey methods in replicating survey-based I-O coefficients

Statistics	Methods							
	SLQ	PLQ	WLQ	CILQ	SCILQ	RLQ	FLQ	SDP
CS	296.441	292.116	352.769	1473.083	1920.705	1041.456	397.785	715.447
SLOPE-C	2.185	2.145	2.415	2.611	3.457	2.557	1.665	2.777
R^2-C	0.499	0.501	0.463	0.450	0.453	0.453	0.461	0.491
SLOPE-R	2.112	2.046	2.351	2.234	2.922	2.213	1.132	3.082
R^2-R	0.538	0.538	0.538	0.455	0.436	0.458	0.374	0.538
US	0.024	0.019	0.056	0.066	0.165	0.054	0.037	0.075
SI	0.692	0.694	0.677	0.660	0.641	0.665	0.663	0.668
MWE	0.002	0.002	0.003	0.002	0.004	0.002	-0.003	0.004
STPE	124.972	122.105	144.495	151.525	190.048	146.904	103.918	158.402

Source: Author's elaboration

Tab. 3 – Ranking of methods on the basis of performances in replicating survey-based I-O coefficients

Statistics	Methods							
	SLQ	PLQ	WLQ	CILQ	SCILQ	RLQ	FLQ	SDP
CS	2	1	3	7	8	6	4	5
SLOPE-C	3	2	4	6	8	5	1	7
R^2-C	2	1	4	8	7	6	5	3
SLOPE-R	3	2	6	5	7	4	1	8
R^2-R	1	2	2	4	5	3	6	1
US	2	1	5	6	8	4	3	7
SI	2	1	3	7	8	5	6	4
MWE	2	1	6	4	7	3	5	8
STPE	3	2	4	6	8	5	1	7
Average rank	2.2	1.4	4.1	5.9	7.3	4.6	3.6	5.6

Source: Author's elaboration

Between the two extremes, there can be found in the order: SLQ (2.2), FLQ (3.6), WLQ (4.1), RLQ (4.6), SDP (5.6) and CILQ (5.9). Thus, of the methods analysed, PLQ perform better in a partitive sense whereas SCILQ performs worse. FLQ, introduced to overcome some limitations related to traditional location quotients, only ranks on the third position and is overtaken by methods such as SLQ and PLQ. However, results refer to application of FLQ with a value of 0.3 assigned to the parameter δ.[13]

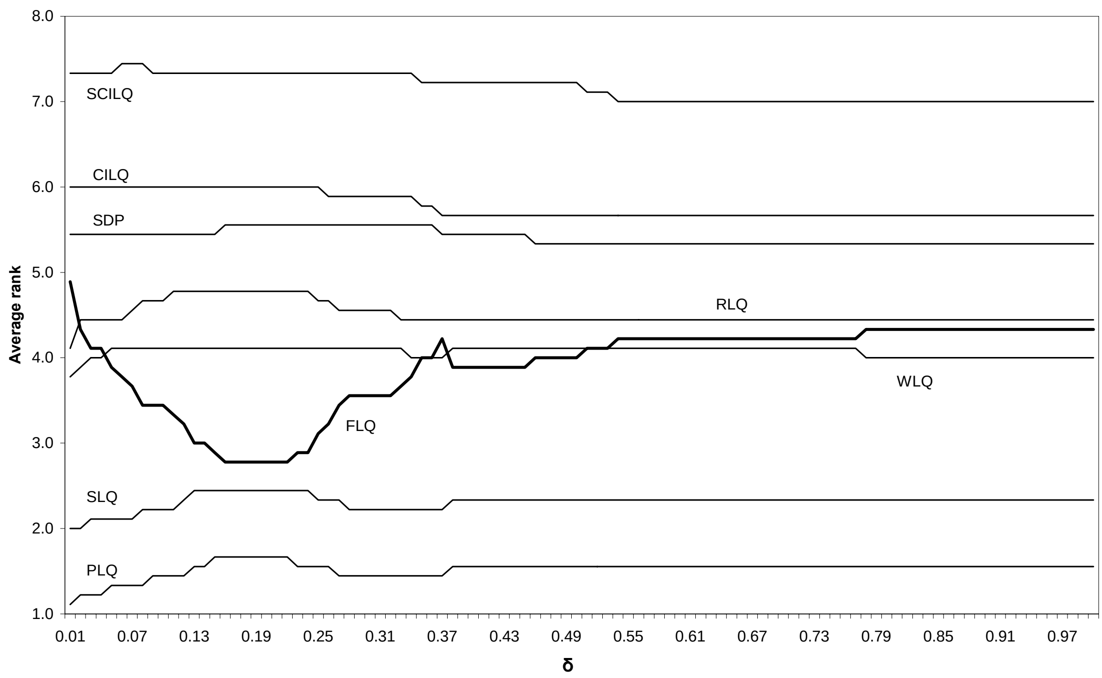
Therefore, it can be interesting to verify the possibility of improving the results produced by FLQ attributing a different value to the parameter δ.

[13] The FLQ's formula (Flegg and Webber, 1997) is: $FLQ_{ij} = \left[\left(E_i^R / E_j^R\right) / \left(E_i^N / E_j^N\right)\right] \cdot \lambda^*$; where E is employment, R is region, N is nation, $\lambda^* = \left[\log_2\left(1 + E^R / E^N\right)\right]^\delta$, $0 \le \delta < 1$, $0 \le \lambda^* \le 1$. The parameter δ should be estimated. On the basis of studies concerning the small English town of Peterborough in 1968 (Morrison and Smith, 1974) and Scotland in 1989 (Flegg and Webber, 1996a, 1996b), Flegg and Webber (1997) argue that an approximate value for δ of 0.3 allows deriving closer estimates to those obtained by surveys than estimates derived by conventional location quotients (SLQ and CILQ). However, the authors remind that more empirical studies are needed to confirm the value of δ.

To explore this hypothesis, a performance analysis, applied to the non-survey methods considered, was repeated 100 times assigning values from 0 to 0.99, with an increment by 0.01, to the FLQ's parameter δ.

Fig. 1 shows the trend of average rank obtained by the methods in replicating survey-based I-O coefficients, attributing increasing values to the FLQ's parameter. As emerges from the figure, the FLQ's average rank function decreases until a certain δ's interval, then starts increasing until stabilising to a value of about 4. For $\delta \leq 0.4$, FLQ would seem to be sensitive to variations of the parameter whereas for bigger values, FLQ would become invariant.

Fig. 1 – Trend of FLQ average rank in reproducing survey-based I-O coefficients matrix



Source: Author's elaboration

For δ's values included in the interval 0.08-0.27, FLQ produces better results than those obtained with $\delta = 0.3$ whereas its optimal values are reached in correspondence of the δ's interval 0.16-0.22.

As far as capability to reproduce survey-based output multipliers is concerned (Tab. 4), firstly, it results that methods would seem to produce better results in comparison to those obtained in terms of I-O coefficients. This is confirmed by improvements highlighted by 6 statistics such as all the regression measures, US and STPE. Even CS improves but only for the methods: CILQ, SCILQ, RLQ and FLQ. On the contrary, SI and MWE worsen for all methods.

Tab. 4 – Performances of non-survey methods in replicating survey-based output multipliers

Statistics	Methods							
	SLQ	PLQ	WLQ	CILQ	SCILQ	RLQ	FLQ	SDP
CS	323.560	312.137	450.827	704.874	1818.724	627.099	292.911	923.661
SLOPE-C	1.021	1.019	1.029	1.044	1.045	1.037	0.991	1.035
R^2-C	0.984	0.984	0.980	0.980	0.964	0.981	0.988	0.979
SLOPE-R	1.022	1.019	1.031	1.046	1.046	1.038	0.990	1.037
R^2-R	0.985	0.986	0.980	0.981	0.968	0.982	0.988	0.978
US	0.002	0.001	0.010	0.022	0.028	0.014	0.022	0.014
SI	0.510	0.516	0.479	0.472	0.421	0.479	0.554	0.427
MWE	0.006	0.005	0.009	0.008	0.011	0.007	-0.004	0.012
STPE	24.938	23.955	30.042	31.239	42.064	29.804	17.195	34.899

Source: Author’s elaboration

Secondly, the method which gets the highest average rank is FLQ (1.7), followed by PLQ (1.9), SLQ (2.9), RLQ (4.6), WLQ (4.7), CILQ (6.0), SDP (6.4) and SCILQ (7.9) (Tab. 5).

Accordingly, FLQ reveals itself to be the best non-survey method to replicate survey-based output multipliers. Superiority of FLQ, in terms of ranking, is confirmed by 8 out of the 9 statistics used. Other good methods are PLQ and SLQ whereas the others produce worse results.

Also in this case, it can be appealing to verify the existence of a different value for the parameter δ which permits to improve FLQ’s performances.

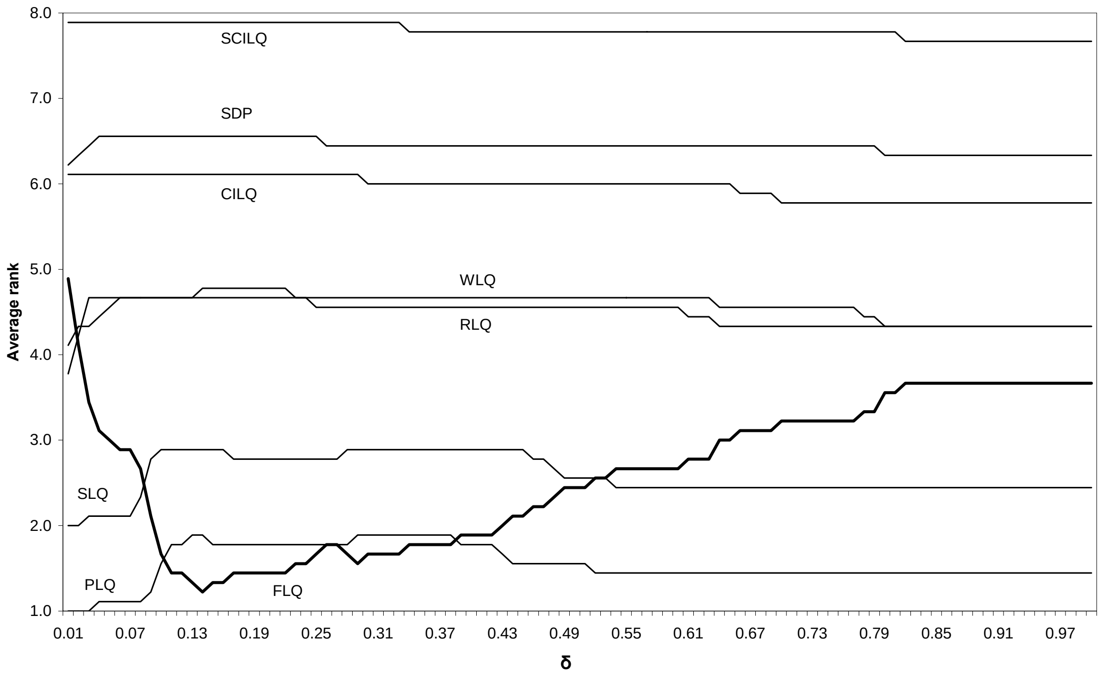
Fig. 2 shows the trend of average rank of methods in reproducing survey-based output multipliers, attributing increasing values to the FLQ’s parameter δ.

Tab. 5 – Ranking of methods on the basis of performances in replicating survey-based output multipliers

Statistics	Methods							
	SLQ	PLQ	WLQ	CILQ	SCILQ	RLQ	FLQ	SDP
CS	3	2	4	6	8	5	1	7
SLOPE-C	3	2	4	7	8	6	1	5
R^2-C	3	2	5	6	8	4	1	7
SLOPE-R	3	2	4	7	8	6	1	5
R^2-R	3	2	6	5	8	4	1	7
US	2	1	3	6	8	4	7	5
SI	3	2	5	6	8	4	1	7
MWE	3	2	6	5	7	4	1	8
STPE	3	2	5	6	8	4	1	7
Average rank	2.9	1.9	4.7	6.0	7.9	4.6	1.7	6.4

Source: Author's elaboration

Fig. 2 – Trend of FLQ average rank in reproducing survey-based I-O multipliers matrix



Source: Author's elaboration

In comparison to the analysis of I-O coefficients, the FLQ's average rank function decreases more rapidly until a certain δ's value, then starts increasing more gradually until reaching a value of about 4.0. Accordingly, the sensitivity of FLQ to the parameter would seem to be bigger. For δ's value included in the interval 0.11-0.24, FLQ's performances can be improved further whereas the optimum is reached in correspondence of a δ's value of 0.14.

4 Concluding remarks

This research has attempted to evaluate performances of eight non-survey methods in reproducing a survey-based I-O model in a both partitive and holistic sense. The survey-based matrix used as a benchmark was the 44-sector 1974 survey-based I-O table constructed for the Italian Marche Region.

The following non-survey methods were explored: the Simple Location Quotient (SLQ), the Purchases-only Location Quotient (PLQ), the West's Location Quotient (WLQ), the Cross-Industry Location Quotient (CILQ), the Semilogarithmic Location Quotient (RLQ), the Symmetric Cross Industry Location Quotient (SCILQ), the Flegg *et al.* Location Quotient (FLQ) and the Supply Demand Pool technique (SDP).

In order to evaluate methods, a package of statistics was selected and applied. This package had to be as wide as possible to diminish the risk that conclusions could be influenced by the statistic used and net of redundant measures which could have had the effect of distorting conclusions on performances of methods. Towards this aim, 35 statistics found in the literature were first explored. Then, they were compared to each other by a correlation analysis applied to a set of data whose variables were the statistics examined and the observations were results of application of statistics to random benchmark matrices compared to random matrices simulating indirectly constructed matrices. On the basis of specific choice criteria, 26 statistics were excluded from the analysis of performances of indirect techniques and the remaining 9 statistics were included in the package of statistics used. These latter were five traditional statistics (CS, squared correlation coefficient applied to both rows, R^2-R, and columns, R^2-C; slope applied to both rows, slope-R, and columns, slope-C) and four general distance statistics (variance of Theil's index of inequality, US; similarity index, SI; mean weighted error, MWE; standardized total percentage error, STPE).

The main results of the performances analysis are as follows.

In general, the non-survey methods examined produce estimates which would seem to be too far from the "true" ones. For this reason, methods should not be used alone but integrated with all available exogenous information within hybrid procedures.

Secondly, methods reveal to be better in replicating survey-based output multipliers rather than I-O coefficients, although not all statistics provide concordant conclusions.

Thirdly, if one had to employ one non-survey method to replicate a survey-based model in both senses (partitive and holistic), the choice should fall on PLQ, which ranks on the first and on the second position, at levels of I-O coefficients and multipliers, respectively. The fact that this method demonstrated to produce better results presents some advantages. First, data requirement of PLQ is smaller than other methods (i.e. WLQ and SDP) and, second, it is not requested estimation of any parameter (like FLQ). However, if the aim was only oriented to conduct an impact analysis, regardless precision of I-O coefficients, FLQ (with $\delta = 0.3$) should be preferred since it demonstrated to overtake all the other methods.

Lastly, following suggestions from a Flegg and Webber's (1997) study to make further regional experiments about estimation of FLQ's parameter δ, FLQ's behaviour was investigated. Results showed that FLQ roughly produces better performances in replicating survey-based I-

O coefficients and multipliers for δ's values between 0.1 and 0.2. Nevertheless, a value of 0.3 demonstrated to produce satisfactory results in reproducing survey-based output multipliers.

However, all the considerations made have to be taken with caution. Firstly, the package of statistics used to make comparisons could not to be exhaustive. Some important statistics could have been neglected and this exclusion might affect results. Secondly, the statistics used could not to be the best measures to compare non-survey methods with the survey-based ones, given the peculiar features characterising I-O tables. Other measures could have given more reliable results. Therefore, an analysis of statistic properties would be necessary.

In any case, made these due specifications, we feel that results of this analysis offer some interesting hints and are encouraging for development of further research in this direction.

References

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

Butterfield, M. and Mules, T. (1980). A Testing routine for evaluating cell by cell accuracy in short-cut regional input-output tables. *Journal of Regional Science* 20: 293-310.

Canning, P. and Wang, Z. (2004). A Flexible Modeling Framework to Estimate Interregional Trade Patterns and Input-Output Accounts. World Bank Policy Research Working Paper 3359, The World Bank Group.

Czamanski, S. and Malizia, E. E. (1969). Applicability and limitations in the use of national input-output tables for regional studies. *Regional Science Association Papers and Proceedings* 23: 65-77.

Dewhurst, J. H. L. (1992). Using the RAS Technique as a Test of Hybrid Methods of Regional Input-Output Table Updating. *Regional Studies* 26: 81-91.

Eskelinen, H. and Suorsa, M. (1980). A note on estimating interindustry flows. *Journal of Regional Science* 20: 261-266.

Flegg, T. A. and Webber, C.D. (1996a). Using location quotients to estimate regional input-output coefficients and multipliers. *Local Economy Quarterly* 4: 58-86.

Flegg, T. A. and Webber, C.D. (1996b). The FLQ formula for generating regional input-output tables: an application and reformation. *Working Papers in Economics* 17, University of the West of England. Bristol: University of the West of England.

Flegg, T. A. and 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. and Webber, C.D. (2000). Regional Size, Regional Specialization and the FLQ Formula. *Regional Studies* 34: 563-69.

Flegg, T. A., Webber, C.D. and Elliot, M. V. (1995). On the appropriate use of location quotients in generating regional input-output tables. *Regional Studies* 29: 547-561.

Gerking, S., Isserman, A., Hamilton, W., Pickton, T., Smirnov, O. and Sorenson, D. (2001). Anti-Suppressants and the Creation and Use of Non-Survey Regional Input-Output Models. In Lahr, M. L. and Miller, R. E. (ed.), *Regional Science Perspectives in Economic Analysis.* Amsterdam: North-Holland, 379-406.

Gilchrist, D. A. and St Louis, L. V. (1999). Completing Input-Output Tables Using Partial Information, with an Application to Canadian Data. *Economic Systems Research* 11: 185-193.

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. and McNicoll, I. H. (1980). A comparison of regional and national technical structures. *Economic Journal* 90: 795-810.

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

ISTAT (1971). *5º Censimento generale dell'industria e dei servizi.* Roma: ISTAT.

ISTAT (1971). *11° Censimento della popolazione e delle abitazioni.* Roma: ISTAT.

ISTAT (1977). Occupati per attività economica e regione, 1970-1976. *Collana di informazione* 3. Roma: ISTAT.

ISTAT (1977). Tavola intersettoriale dell'economia italiana per l'anno 1974. *Collana d'informazione* 5. Roma: ISTAT.

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

Jackson, R.W. and Murray, A. (2004). Alternative Input-Output Matrix Updating Formulations. *Economic Systems Research*, 16: 135-148.

Jalili, A.R. (2000). Evaluating Relative Performances of Four Non-Survey Techniques of Updating Input-Output Coefficients. *Economics of Planning* 33: 221-237.

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

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. and Karunarante, N.D. (1979). *Regional Economic Planning: Generation of Regional Input-Output Analysis.* London: Croom Helm.

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 Lahr M. L. and Dietzenbacher (ed.), *Input-Output Analysis: Frontiers and Extensions.* London: Palgrave, 1-31.

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

Madsen, B. and Jensen-Butler, C. (1999). Make and Use Approaches to Regional and Interregional Accounts and Models. *Economic Systems Research* 11: 277-299.

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

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

Mogorovich, L. (1987). *Tavole input-output regionali: tecniche indirette e sensibilità agli errori.* Dissertation in Statistics, University of Trieste, Italy.

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

Okuyama, Y., Hewings, G.J.D., Sonis, M. and Israilevich, P.R. (2000). An Econometric Analysis of Biproportional Properties in an Econometric Input-Output Modeling System. Discussion Papers, The Regional Economics Applications Laboratory (REAL). South Matthews, Urbana: REAL.

Oude Wansink, M. J. and Maks, J. A. H. (1998). Constructing Regional Input-Output Tables. Working Paper, Maastricht University.

Richardson, H. W. (1985). Input-Output and Economic Base Multipliers: Looking Backward and Forward. *Journal of Regional Science* 25: 607-771.

Robison, S., Cattaneo, A. and El-Said, M. (2001). Updating and Estimating a Social Accounting Matrix Using Cross Entropy Methods. *Economic Systems Research* 13: 47-64.

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

Santeusanio, A. (1979). La tavola economica intersettoriale delle Marche – 1974. In Santeusanio, A., Mattioli, E. and Turci, M. C., *Primi studi sulla struttura della economia delle Marche*. Faculty of Economics and Commerce of the University of Urbino, Institute of Mathematics and Statistics (IMS). Ancona, Italy: IMS, 30-49.

Sawyer, C. and Miller, R. (1983). Experiments in regionalization of a national input-output table. *Environment and Planning A* 15: 1501-1520.

Schaffer, W. and Chu, K. (1969a). Nonsurvey Techniques for Constructing Regional Interindustry Models. *Papers and Proceedings of the Regional Science Association* 23: 83-101.

Schaffer, W. and Chu, K. (1969b). Simulating Regional Interindustry Models for Western States. A Program on Regional Industrial Development, Discussion Paper 14, Georgia Institute of Technology, Georgia.

Stevens, B. H. and Trainer, G. A. (1976). The Generation of Error in Regional Input-Output Impact Models. Working Paper No. A1-76, Regional Science Research Institute (RSRI). Peace Dale, Rhode Island: RSRI.

Strassoldo, M. (1988). *Contabilità macroeconomica e tavole input-output regionali*. Milan: Giuffrè Editore.

Theil, H. (1971). *Applied Economic Forecasting*. Amsterdam: North-Holland.

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

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

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

Willis, K. G. (1987). Spatially Disaggregated Input-Output Tables: An Evaluation and Comparison of Survey and Nonsurvey Results. *Environment and Planning A* 19: 107-116.