The S-RAS Method and Its Application to Sector Disaggregation of Input-Output Tables

Authors: Liu Xinjian, Yang Qian, Song Zhongwei, Liu Xinjian, Yang Xi

Date: 2022-10-12T00:00:00+00:00

Abstract

The RAS method is one of the most important techniques for compiling and revising input-output tables. Building upon the basic RAS method and existing improved RAS methods, this paper proposes the S-RAS method and its operational procedures. The S-RAS method comprises four major steps: matrix preprocessing, applicability adjustment, obtaining the target matrix using the basic RAS method, and new table formation. The S-RAS method extends the RAS method from merely updating the intermediate flow matrix to the entire input-output table, wherein the applicability adjustment step enhances the necessary treatment of implicit conditions in the application of the RAS method. The S-RAS method was developed during a project tasked with sector disaggregation; consequently, the paper demonstrates its applicability through two sector disaggregation examples.

Full Text

S-RAS Method and Its Application in Sector Division of Input-Output Tables

Liu Xinjian, Yang Qian, Song Zhongwei
(School of Economics and Management, Yanshan University, Qinhuangdao 066004, China)

Abstract

The RAS method is one of the most important techniques for compiling and revising input-output tables. Building upon the basic RAS method and existing improved RAS methods, this paper proposes the S-RAS method and its operational procedures. The S-RAS method comprises four major steps: matrix preprocessing, applicability adjustment, obtaining the target matrix using the basic RAS method basic RAS method, and new table formation. It extends the application of the RAS method from merely updating the intermediate flow matrix to the entire input-output table, with the applicability adjustment step supplementing the necessary treatment of implicit conditions in RAS method applications. The S-RAS method was developed during a research project on sector division tasks; therefore, this paper demonstrates its applicability through two examples of sector division.

Keywords: input-output; RAS method; S-RAS method; sector splitting

Author Biographies: Liu Xinjian (1963-), male, from Jishan, Shanxi, Ph.D. in Management Engineering from the Institute of Systems Science, Chinese Academy of Sciences, Professor at the School of Economics and Management, Yanshan University, Research interests: input-output analysis, E-mail: lxj6309@126.com; Corresponding Author: Yang Qian (1995-), female, from Xinji, Hebei, Graduate student at the School of Economics and Management, Yanshan University, Research interests: input-output analysis, Contact: 17743771695, E-mail: 784355748@qq.com; Song Zhongwei (1994-), male, from Handan, Hebei, Graduate student at the School of Economics and Management, Yanshan University, Research interests: input-output analysis, E-mail: 1942673482@qq.com.

Funding: This paper is one of the research outcomes of the 2017 Input-Output Research Project of the National Bureau of Statistics, "Research on the Impact of Structural Changes in Final Use on Economic Development."

Introduction

Input-output tables form the foundation for applying input-output techniques to practical economic analysis. However, compiling an input-output table with hundreds of sectors requires substantial investment of human, material, and financial resources, making it impossible to produce continuous annual time series. The continuous annual time series tables that do exist are all extended from basic tables using limited routine statistical data. Among various extension table compilation methods, the RAS technique stands as one of the most important. China compiles a special survey-based table every five years, with an extended table compiled in between based on routine statistical data and limited surveys. The RAS method also serves as the fundamental technique for extended table compilation.

The RAS method, also known as the biproportional scaling method, is a type of scaling algorithm in matrix transformation techniques. Biproportional matrix adjustment techniques were first proposed in the 1930s, also referred to as "iterative proportional fitting" or "raking" [1]. Stone introduced the RAS method into input-output techniques in 1960 [2], where it gained widespread application as an auxiliary technique for compiling input-output tables. The RAS method can reduce a problem with n^2 unknowns to one with only $2n$ unknowns. Since the 1990s, scholars have demonstrated its rationality from the perspectives of instrumental variable regression and optimization models [3,4].

Due to the extremely limited data information relied upon by the RAS method, its accuracy is necessarily constrained. Consequently, researchers have continuously proposed improved RAS methods under conditions of additional information. First, the improved RAS method was proposed utilizing the zero-preservation property [5], followed by the TRAS method that makes full use of information on the sums of known submatrix elements [6]. For cases where matrix elements become negative, Oosterhaven (2005) [7] proposed the GRAS (Generalized RAS) method, primarily applied to Social Accounting Matrix (SAM) updating. Several improved algorithms have been built upon GRAS: Chen Ronghu [8] improved the objective function of the GRAS method and proposed the AGRAS method using an adaptive search approach to find its optimal solution; Temursho and Oosterhaven [9] proposed the MA-RAS method extending GRAS techniques to multi-regional or transnational environments; Lenzen et al. [10] proposed the KRAS method capable of balancing input-output tables and social accounting matrices under conflicting external information or inconsistent constraints; Zhou Nannan and Li Baoyu [11] proposed the DRAS method for the joint balancing and forecasting of two matrices, applicable to fund flow matrix extension table estimation; Holý and Šafr [12] proposed a multidimensional RAS method for disaggregating annual national tables into more detailed tables (regional, quarterly, domestic/import) to ensure consistency between classified and aggregate tables. Some scholars have verified through practical applications that the RAS method or revised RAS methods demonstrate good estimation performance [13,14], which explains their widespread adoption in input-output table updating.

Throughout the compilation history of a country' s input-output tables, sector classification criteria frequently change with the continuous development of statistical systems and techniques. When examining China' s historical input-output tables compiled by the National Bureau of Statistics, we observe increasingly refined sector classifications. Consequently, to obtain consistent sector classifications, the task of splitting past coarse sectors into finer ones emerges. In addressing this problem, we identified certain detailed issues in traditional RAS method applications and their extensibility, leading to the proposal of the S-RAS method, or Systematic RAS method, which we apply to sector division tasks. This paper first elaborates on the algorithmic principles and application steps of the RAS method and improved RAS methods, then presents the S-RAS method, and applies it to the division of the post and telecommunications sector in China' s 1992 input-output table and the textile industry sector in China' s 2018 input-output table.

2 RAS Method and Its Improved Model Principles and Steps

The basic RAS method serves as the foundation for various improved RAS methods, with its algorithm forming the main axis of all improved RAS algorithms. This section introduces the principles, implicit conditions, mathematical properties, basic algorithmic steps of the basic RAS method, and the specific procedures of improved RAS methods. Here, we explicitly clarify two implicit conditions for RAS method application.

2.1 Principles and Basic Steps of the Basic RAS Method

The basic data for the RAS method consists of either: (1) the intermediate flow matrix of a base input-output table as the starting point, along with the row sum vector and column sum vector of the intermediate flow matrix of the target input-output table; or (2) the direct input coefficient matrix corresponding to the base table, the final use matrix of the target input-output table, the primary input matrix, and the total output vector.

(1) Basic RAS Method Principle [4,15]

Let X be the known intermediate flow matrix of the base table, and U and V be the row sum vector and column sum vector of the target table, respectively. If there exist vectors R and S such that:

$$\hat{R}Xe = U \quad \text{and} \quad e\hat{R}XS = V$$

hold true, then $\hat{R}XS$ is taken as the intermediate flow matrix of the target table. Here, e denotes a vector of all ones, called the summation vector, whose row/column attribute and dimension depend on the vector being summed; a hat symbol ($\hat{}$) over a vector symbol indicates a diagonal matrix with the vector' s elements as its main diagonal elements. If X is unknown but the base table' s input coefficient matrix A and the target table' s total output Q are known, then let $X_0 = A\hat{Q}$, and replace X with X_0 in the equation. Thus, the input coefficient matrices of the new and base tables have the relationship $A_1 = \hat{R}A\hat{S}$.

The model has two implicit conditions:

Condition 1: $e^T U = Ve$, meaning the sum of all elements of the row sum column vector U equals the sum of all elements of the column sum row vector V; otherwise, iteration cannot converge. This condition is called the balance requirement.

Condition 2: Since zero cannot be a denominator when calculating row (column) multipliers, no element in the target table' s row sums (column sums) can be zero, and no row (column) of the base table can have all zero elements; otherwise, iteration cannot proceed smoothly. This is called the non-zero requirement.

The RAS method has the following important mathematical properties:

Property 1: Mathematically, the solution to equation system (1) is not unique. If (R, S) is a solution, then given any real number $\alpha \neq 0$, $(\alpha R, \alpha^{-1} S)$ remains a solution. However, $\hat{R}XS$ is unique, meaning the RAS method' s solution for the intermediate flow matrix is unique.

Property 2: Intermediate flow cells or direct input coefficients that are zero in the base table remain zero in the updated target table. This property is called the zero-preservation property of the RAS method and serves as the theoretical basis for the key steps of improved RAS methods.

(2) Basic RAS Method Algorithmic Steps

The essence of the RAS method is solving equation system (1), which constitutes a nonlinear system without analytical solutions; therefore, iterative methods are used in practice.

Assume after k iterations, matrix X_k is obtained such that $eX_k = V$ holds. Let δ be an arbitrarily small real number reflecting the desired precision. The RAS method' s solution steps are as follows:

Step 1: Compute $X_{k+1} = \hat{R}_{k+1} X_k$. If $\|U - X_{k+1} e\|_F < \delta$ (where $\|\cdot\|_F$ denotes the matrix Frobenius norm), then stop iteration and let $X = X_{k+1}$; otherwise, proceed to the next step.

Step 2: Compute row scaling factors: $\hat{R}_{k+1} = \hat{U}/(X_k e)$.

Step 3: Compute column scaling factors: $\hat{S}_{k+1} = \hat{V}/(e\hat{R}_{k+1} X_k)$. If $\|V - eX_{k+1}\|_F \leq \delta$, then stop iteration and let $X = X_{k+1}$; otherwise, proceed to the next step.

Step 4: Let $k = k + 1$ and repeat Steps 1-3.

2.2 Improved RAS Method

The data foundation of the improved RAS method is that, in addition to the basic data required by the basic RAS method, some cells in the target table' s intermediate flow matrix have relatively reliable information. Therefore, these cells' values need not be adjusted in the RAS method.

Based on this, the improved RAS method utilizes the zero-preservation property of the basic RAS method by setting known data cells to zero, subtracting the values of known cells from the corresponding row and column sums, then executing the basic RAS method, and finally restoring the known values to their positions to obtain the target table [5].

3 S-RAS Method and General Steps in Input-Output Table Revision Applications

The S-RAS method represents a generalized extension of the basic RAS method, expanding from the intermediate flow matrix to the entire input-output table.

This section first presents the general operational steps of the S-RAS method, followed by special processing for sector division problems.

3.1 Principles of the S-RAS Method

As a biproportional matrix adjustment technique, the fundamental application condition of the RAS method is a base matrix and an unknown target matrix with known row and column sums. The basic RAS method assumes a known intermediate flow matrix as the base matrix; for the target intermediate flow matrix to be estimated, the column sums of total intermediate use and row sums of total intermediate input are known. If the currently blank fourth quadrant is filled with zeros, the newly formed input-output table constitutes a complete matrix, which can evidently also be updated using the RAS method for matrix element updating.

For a matrix composed of the complete input-output table (hereinafter referred to as the full-table matrix), assuming the original number of production sectors is n, if the final use column Y and primary input row Z are known, we can construct an $(n+1)$-order full-table matrix where the $(n+1)$-th row and $(n+1)$-th column are set to zero, then subtract Y from the total output column and Z from the total input row, thereby obtaining a new full-table matrix for target table estimation using the basic RAS method. Of course, this situation could also directly apply the basic RAS method to the intermediate flow matrix.

If certain elements or sums of elements in the target table are known (these elements may belong to the second or third quadrants), then the improved RAS method can be used for target table estimation.

This paper designates the above RAS method for full-table matrices, supplemented with processing that considers some implicit conditions, as the S-RAS method—signifying a more systematic RAS method for input-output table updating.

3.2 General Operational Steps of the S-RAS Method

The prerequisite for applying the S-RAS method is knowing a full-table base matrix, and the "row sum" column and "column sum" row of a full-table target matrix. If no constraint information exists for elements of the full-table target matrix, then for a standard input-output table, the elements of the "row sum" column and "column sum" row equal the corresponding elements of the total output column. If certain elements or other constraint information are known, then matrix preprocessing is required, which will cause corresponding elements of the "row sum" column and "column sum" row to differ.

Step 1: Matrix Preprocessing

In S-RAS method applications for input-output table updating, three table concepts need definition. First is the base table—a fully known input-output table; second is the target table—the desired input-output table to be obtained; third is the benchmark table—the input-output table after standardized processing of the base table. When applying the basic RAS method, the benchmark table and base table should share the same structure as the target table, necessitating some essential preprocessing of the benchmark and base tables.

Step 1.1: Establish coordinated and consistent structures for base, benchmark, and target tables. The sector count and arrangement of the base input-output table used as the starting point may differ from the target table. To facilitate programmatic data processing, adjustments should be made to maintain basic consistency between them. The benchmark and target tables should have completely identical structures in terms of sector count, sector ordering, final use item arrangement, and primary input item arrangement. The structural adjustment requirements are typically determined by the research and analysis task.

Step 1.2: Define the benchmark matrix. The so-called benchmark matrix is formed by filling the fourth quadrant of the benchmark table with zeros to create a full-table matrix, designated as the benchmark matrix. Denoted as:

$$X_0 = \begin{pmatrix} Z & F \\ P & 0 \end{pmatrix}$$

where Z is the original intermediate flow matrix, F is the original final use matrix, and P is the original primary input matrix.

Step 1.3: Determine the target matrix' s row sum U_1 and column sum V_1. The target matrix is the updated matrix derived from the benchmark matrix, i.e., the full-table matrix of the new input-output table. If certain elements in the target table are known, corresponding adjustments should be made to U_1 and V_1.

Step 2: Matrix RAS Method Applicability Adjustment

According to the mathematical principles of the basic RAS method and its algorithmic implicit requirements, to satisfy iteration feasibility and improve accuracy, further processing of the benchmark matrix and target matrix row/column sums is necessary.

Step 2.1: Systematic error handling due to RAS method zero-preservation property. Because of the RAS method' s zero-preservation property, when a row (column) of the benchmark matrix contains all zeros but the corresponding row sum (column sum) in the target matrix is non-zero, the benchmark matrix requires appropriate processing to prevent corresponding elements in the target matrix from also being zero, thereby maintaining the original balance of the base table.

Step 2.2: Processing to satisfy RAS method non-zero requirements. According to the RAS method' s non-zero requirements, on one hand, zero elements in the target matrix' s row sum U_1 and column sum V_1 must be assigned values; on the other hand, when a row (column) of the benchmark matrix contains all zeros, relevant cells in this row (column) must be assigned values. To minimize errors, all assigned values should be sufficiently small.

Step 3: Obtain Revised Matrix Using Basic RAS Method

Based on the applicability-adjusted benchmark matrix X_0, target matrix row sum U_1, and column sum V_1, implement the basic RAS method to obtain the target matrix.

Step 4: New Table Formation

Embed the target matrix into the target table framework, and appropriately process data significant digits (for economies of considerable scale, typically rounding to integers or one decimal place). Finally, adjust unreasonable data according to economic significance to obtain the complete target input-output table.

3.3 Special Processing of S-RAS Method for Sector Division

In sector division tasks for input-output table updating, there are actually two base tables—two fully populated input-output tables: the first is the table before sector division, containing division constraints; the second is the reference benchmark input-output table, i.e., a table from another year or similar economy with subdivided sectors already classified; the benchmark table is the base table after standardized processing; the target table is the table with divided sectors and completed data filling, though most data remain unknown before RAS method implementation except for known data. For sector division tasks, the first base table and target table belong to the same economy and year, with the vast majority of sectors and data being identical, and certain aggregate constraints existing for splitting coarse sectors into fine ones. Therefore, some special processing is required when applying the S-RAS method.

(1) Adjust the sector ordering of the base and benchmark tables so that, except for the specific sector to be processed, they align with the target table structure (including final use and primary input), with the specific sector to be processed positioned adjacent. During processing, some necessary sector merging and data adjustments may be required. Let the base table structure awaiting sector division be as shown in Table 1 , with the target table splitting base table sector a into subdivided sectors i and j.

Table 1. Input-Output Table Structure without Split Sectors [Table content showing sectors 1, a, n and their flows]

Note: To keep the table concise, ellipses indicating omitted data are omitted from the table, the same applies to Table 2 .

(2) **Definition of benchmark matrix.** In sector division tasks, since most data are identical between the first base table and target table, set all elements of the benchmark table to zero except those involving subdivided

sectors i and j, as shown in Table 2. Fill the fourth quadrant cells with zeros to form the benchmark matrix X_0.

Table 2. Benchmark Input-Output Table Structure [Table content showing sectors 1, i, j, n and their flows]

(3) **Target matrix** is the updated matrix derived from the base table and benchmark matrix, containing data for subdivided sectors after division. Estimate the total output of specific subdivided sectors in the target table based on the base table and other known information, i.e., the division. The preliminary target matrix row sum vector U_1 and column sum vector V_1 are:

$$U_1 = [\text{row sums after subtracting known cells}]$$
$$V_1 = [\text{column sums after subtracting known cells}]$$

where $P\Lambda$ represents the values of total output for each sector in the target matrix and row sums of primary input items after subtracting known cell data from corresponding rows in the base table; $F\Lambda$ represents the values of total input for each sector in the target matrix and column sums of final use items after subtracting known cell data from corresponding columns in the base table.

(4) **In matrix RAS method applicability adjustment**, first, row-wise, if the benchmark matrix' s x-th row has zero product flows to subdivided departments (i.e., $z_{xi} = z_{xj} = 0$), but the corresponding row sum in the target matrix $U_x \neq 0$, then assign values to x_{iz} and x_{jz} in the benchmark matrix in a certain manner; the same applies column-wise. Second, determination of intra-sector product flows. The internal product flow of the sector awaiting division in the base table is Z_{aa}; after division, it constitutes intra-sector product flows among subdivided sectors in the target table, represented by submatrix:

$$\begin{pmatrix} z_{ii} & z_{ij} \\ z_{ji} & z_{jj} \end{pmatrix}$$

satisfying the condition $z_{ii} + z_{ij} + z_{ji} + z_{jj} = Z_{aa}$. For these four elements, if additional information is available, it should be determined based on reliable information and then treated as known values; if no further information exists, the TRAS method can be applied, or the proportional relationships from the benchmark table can be used to allocate Z_{aa}.

The above steps split a specific sector into two sectors. If multiple sectors need division, the basic S-RAS method steps can be executed repeatedly, with sequential splitting being more precise. Although theoretically possible to split a sector into more than two subdivisions and the basic S-RAS method imposes no theoretical limit on the number of subdivided sectors, the most reliable approach is to first split a sector into two, then further subdivide as needed.

4 Application of S-RAS Method in Sector Division

The S-RAS method was developed when facing sector division requirements in a research project. Therefore, we illustrate its application through two sector division examples below, one from the original project and another testing against the latest release from the National Bureau of Statistics.

4.1 Application in China' s 1992 Input-Output Table Post and Telecommunications Sector Division

Input-output tables compiled by the National Bureau of Statistics since 1997 have separated post and telecommunications sectors, with post grouped with transportation and warehousing in the corresponding 42-sector tables, and telecommunications included in the information transmission sector. The 1992 table merged post and telecommunications into a single sector called 'Post and Telecommunications,' which does not meet the needs of relevant information industry analysis. Below, using the 1997 input-output table structure as the reference structure for the target table, we apply the S-RAS method to split the 1992 input-output table' s post and telecommunications sector into two subdivided sectors: post and telecommunications.

4.1.1 Matrix Preprocessing Before division, we first preprocess the 1997 and 1992 input-output tables to align the benchmark table with the target table structure and obtain the initial benchmark matrix.

(1) Sector structure adjustment

The 1997 large table has 124 sectors, while the 1992 large table has 118 sectors. Beyond the post and telecommunications difference, the two tables also differ in other sectors. To achieve one-to-one sector correspondence, through comparative adjustments, we align the sectors of the two years (excluding post and telecommunications), with relevant sector mergers shown in Table 3 .

Table 3. Example of Sector Integration of Input-Output Tables in 1997 and 1992 [Table showing sector mergers: coal mining, grain processing, meat processing, etc.]

After merging, both the 1997 and 1992 input-output tables are integrated into 107 sectors including post and telecommunications, with consistent sector ordering between the two years.

(2) Final use matrix and primary input matrix processing

In unified processing, the final use and primary input structure items and ordering of the two years' input-output tables should also be aligned. In the final use section, since 1992 combines imports and exports as net exports, we also combine imports and exports as net exports for 1997. Although both tables have error terms or other items, the error terms across the two years are not comparable. Therefore, we allocate the 1992 error term proportionally based on the absolute values of final use items in the 1992 post and telecommunications sector. For other sectors, to reduce artificial interference factors, we do not allocate error terms. The 1992 and 1997 primary input items are consistent in category and ordering, requiring no adjustment.

(3) Data supplementation

The key to supplementing row and column control totals lies in obtaining the proportional relationship between the production volumes of the two subdivided sectors. In the National Bureau of Statistics' annual statistical data, the 1992 business volume figures for post and telecommunications are given as 6.436 billion yuan and 22.657 billion yuan, respectively.

The total output of the post and telecommunications sector in the 1992 input-output table is 25,414.99 million yuan. Decomposing this total output using the business volume proportions yields total outputs of 5,622.151497 million yuan for the post sector and 19,792.8385 million yuan for the telecommunications sector, which are supplemented into the target table to complete the row and column control totals.

(4) Zero out known cells to form benchmark matrix

Combine the benchmark table' s intermediate flow matrix, final use matrix, and primary input matrix according to the input-output table format to form the benchmark matrix, filling the fourth quadrant elements with zeros. Set all cells of the benchmark matrix to zero except those related to the specific subdivided sectors, obtaining an 111×113 benchmark matrix X_0. Subtract the determined values of known cells outside the subdivided sectors from the 1992 total output and total input of each sector to obtain the preliminary target matrix row sum U_1 and column sum V_1.

4.1.2 Matrix RAS Method Applicability Adjustment According to the implicit conditions for basic RAS method application, to ensure smooth RAS solution and improve accuracy, some applicability adjustments are necessary.

(1) Handling cases where 1997 flows are zero but 1992 flows are non-zero

Comparing row and column data between the 1997 post and telecommunications sectors and the 1992 post and telecommunications sector, row-wise, 26 sectors including planting have zero product flows into post and telecommunications in 1997, but non-zero flows into post and telecommunications in 1992. Due to the basic RAS method' s zero-preservation property, this situation introduces errors, necessitating non-zero treatment of corresponding 1997 zero values. Assume that for a given sector r in 1997, the product values flowing into post and telecommunications sectors satisfy Equation (7):

$$\frac{z_{r,i}}{Q_i} = \frac{z_{r,j}}{Q_j} = \frac{z_{r,u}}{Q_u}$$

where $z_{r,i}$ and $z_{r,j}$ represent the product values flowing from sector r into post and telecommunications sectors in 1997, respectively; Q_i and Q_j represent the total inputs of post and telecommunications sectors; $z_{r,u}$ represents the product value flowing from sector r into the post and telecommunications industry in 1992; Q_u represents the total input of the post and telecommunications industry. Simultaneously, the 1997 $z_{r,i}$ and $z_{r,j}$ must satisfy Equation (8):

$$z_{r,i} + z_{r,j} = z_{r,u}$$

From Equations (7) and (8), we can solve for:

$$z_{r,i} = z_{r,u} \cdot \frac{Q_i}{Q_i + Q_j}, \quad z_{r,j} = z_{r,u} \cdot \frac{Q_j}{Q_i + Q_j}$$

In the final use section, column-wise, inventory increases for both post and telecommunications sectors are zero in 1997, while inventory increase for post and telecommunications is non-zero in 1992. Referencing fully separated post and telecommunications input-output tables from 2017 and 2018, post products are used for both fixed capital formation and inventory increase, while telecommunications products have zero values for both items. Therefore, we add the 1992 inventory increase entirely to the post sector. With the inventory increase values for post and telecommunications sectors in the target table now determined, set the corresponding cells for both sectors' inventory increase in the benchmark matrix to zero, and subtract the determined inventory increase values from the target matrix' s row and column control totals, excluding them from the basic RAS method solution process.

(2) Processing to satisfy RAS method non-zero requirements

Since the RAS method calculation process requires non-zero values for all elements in the target matrix's row and column sums, we assign values to zero cells in 1992 row and column sums. Considering inter-sector proportions and minimizing errors, we change 16 sectors' row sum cells including animal husbandry from 0 to 10^{-9}, and change column sum cells for total social consumption, fixed capital formation, and inventory increase from 0 to 10^{-9}.

Simultaneously, no row or column in the benchmark matrix can have all zero elements. Row-wise, assign values to sectors with zero product flows into post and telecommunications in 1997, changing product values from 15 sectors including animal husbandry from 0 to 10^{-9}. Column-wise, change product values from post and telecommunications sectors flowing into scrap and waste materials, as well as those for total social consumption, fixed capital formation, and inventory increase from 0 to 10^{-9}—values sufficiently small to not affect inter-sector proportional relationships.

(3) Determination of intra-sector product flows for post and telecommunications

The internal product flow within post and telecommunications in 1992 is 356. We now need to determine internal product exchanges between post and telecommunications sectors. Based on the proportional relationships of elements in the submatrix formed by internal product flows between post and telecommunications sectors in 1997:

$$\begin{pmatrix} z_{ii} & z_{ij} \\ z_{ji} & z_{jj} \end{pmatrix}$$

we obtain the corresponding submatrix for 1992. Finally, treat each element of this submatrix as determined values in the benchmark matrix, adjusting the benchmark and target matrix row and column control totals accordingly.

4.1.3 Update Calculation Using Basic RAS Method and New Table Formation Based on the adjusted benchmark matrix X_0, target matrix row sum U_1, and column sum V_1, perform basic RAS method adjustment to obtain the updated target matrix. Embed the three parts of the target matrix—intermediate flows, final use, and primary input—into the target table framework, round data related to post and telecommunications sectors, then fill in other 1992 sector data and subdivided sector known values into the target table to obtain the complete input-output table.

Through computational verification, the 1992 input-output table with separated post and telecommunications sectors satisfies balance requirements and remains consistent with known information constraints. Compared with separately applying RAS method adjustments to the three parts of the input-output table (intermediate flow matrix, final use matrix, and primary input matrix), which requires additional artificial setting of total primary input and total final use for subdivided sectors, the S-RAS method requires fewer artificially set values, thereby reducing systematic errors.

4.2 Application of S-RAS Method in China's 2018 Input-Output Table Textile Industry Division

To evaluate the accuracy of the S-RAS method in input-output table division, this section conducts a division test on the 2018 input-output table with known classification results and compares accuracy with traditional proportional division method results.

4.2.1 Division of 2018 Textile Industry Using S-RAS Method The division task involves using the S-RAS method to split the textile industry (merged from textiles and textile clothing, footwear, hats, leather, down, and related products) back into two sectors: textiles and textile clothing, footwear, hats, leather, down, and related products. In operation, the 2017 input-output table serves as the benchmark table, with the target table structure consistent with the benchmark table.

(1) Matrix preprocessing

The 2017 and 2018 input-output tables are basically consistent in sector classification. To examine the feasibility of the S-RAS method when dividing large-scale input-output tables, we process the sector classifications of both tables. First, both tables are processed into 42-sector tables with consistent ordering for storage and backup. Then, the 2018 textiles and textile clothing, footwear, hats, leather, down, and related products sectors are merged into a single textile industry sector, and the row and column sum data required for the S-RAS method are obtained.

In the final use section, imports and exports for both 2017 and 2018 are combined as net exports. Combine the benchmark table' s intermediate flow matrix, final use matrix, and primary input matrix according to the input-output table format to form the benchmark matrix, filling the fourth quadrant elements with zeros. Set all cells of the benchmark matrix to zero except those related to the specific subdivided sectors, obtaining a 46×48 benchmark matrix X_0. Subtract the determined values of known cells outside the subdivided sectors from the 2018 table' s sector total outputs and total inputs to obtain preliminary target matrix row sum U_1 and column sum V_1.

(2) Applicability adjustment

Regarding systematic errors caused by the RAS method' s zero-preservation property, comparison reveals no cases where flows between a sector and the textile industry are zero in 2017 but non-zero in 2018, so no assignment processing is needed. According to non-zero requirements, row-wise, change product values flowing from petroleum and natural gas extraction products and metal ore mining products into textiles and textile clothing, footwear, hats, leather, down, and related products sectors from 0 to 10^{-9}. Column-wise, change product values flowing from textiles and textile clothing, footwear, hats, leather, down, and related products sectors into government consumption expenditure and total fixed capital formation from 0 to 10^{-9}. All elements in the target matrix' s row and column sums are non-zero.

The 2018 intra-sector intermediate flow in the textile industry is 334,855,057.8 billion yuan. Based on the proportional relationships of elements in the submatrix formed by product flows between the two subdivided sectors in 2017:

$$\begin{pmatrix} 142026362 & 132874502 \\ 9187354 & 52427144 \end{pmatrix}$$

we obtain the corresponding submatrix for 2018. Then treat each element value of this submatrix as determined values in the benchmark matrix, adjusting the benchmark and target matrix row and column control totals accordingly.

**(3) Based on the adjusted benchmark matrix X_0, target matrix row sum U_1, and column sum V_1, perform basic RAS method adjustment to obtain the updated target matrix. Substitute the relevant sector data for textiles and textile clothing, footwear, hats, leather, down, and related products from the target matrix into the 2018 42-sector table to obtain the complete 2018 target input-output table.

4.2.2 Error Analysis Many researchers have evaluated errors in the RAS method. Theoretically, this input-output table revision method has no strict error control capability, only achieving so-called minimum cross-entropy minimization [15], making it closer to the benchmark matrix in an entropy sense. If the benchmark and target tables are very similar, such as adjacent years, proportional decomposition may be more accurate than the RAS method. To verify this hypothesis, we calculate and compare both methods for the 2018 textile industry sector division problem described above.

(1) Splitting 2018 textile industry using traditional proportional method

The so-called traditional proportional method splits row and column data of the combined sector in the target table according to the corresponding proportions of the two subdivided sectors in the benchmark table. Building upon the data preprocessing for S-RAS method division, we calculate the proportions of the two subdivided sectors relative to their sum in both row and column directions from the 2017 table, then use these proportions to allocate flows in the 2018 textile industry sector.

The 2018 intra-sector product flows in the textile industry are decomposed using the same processing method as in the S-RAS method described earlier. Obtain product flows for textiles and textile clothing, footwear, hats, leather, down, and related products sectors in both row and column directions, substitute them into the 2018 42-sector input-output table to obtain the target input-output table split using the traditional proportional method.

(2) Error comparison

Error evaluation for updating methods typically involves synthesizing differences between the matrix obtained by the updating method and the corresponding matrix in the actual table, either by calculating the average of absolute values of elements in the difference matrix or by calculating the mean squared error. For full-table updating, all elements of the full-table matrix must be calculated, whereas sector division only involves partial elements. For a full-table matrix with n production sectors, six final use items, and four primary input items, a one-into-two division task involves $4(n+4)$ elements; thus, error analysis only calculates the average error of these elements.

Based on the 2018 input-output table textile industry division results, the comprehensive mean squared error of proportional method division is 1,018,859.28, while that of the S-RAS method is 1,391,033.71, with the former being significantly smaller than the latter. Of course, this error calculation is a black-box operation that completely ignores whether the National Bureau of Statistics made any mutual reference when compiling the input-output tables for the two years.

Conclusion

The key to the basic RAS method and its improvements lies in fully applying known information to improve target table estimation accuracy. The S-RAS method extends the RAS method used for intermediate flow matrix updating to the full-table matrix, considering both the semantic requirements as economic data and the rule requirements for mathematical operations. This approach enhances the capability of the RAS method for input-output table updating and remains compatible with traditional improved RAS methods. In implementing sector division tasks, the target table only requires known subsector total outputs, without needing to obtain subdivided sector total primary input and total final use, reducing statistical error risk. The applicability adjustment in Step 2 of the S-RAS method supplements the necessary treatment of implicit conditions in RAS method applications, improving the rationality of RAS updating results.

References

[1] Ireland C T, Kullback S. Contingency Tables with Given Marginals[J]. Biometrika, 1968, 55(1): 179-188.

[2] Lahr M, Mesnard L D. Biproportional Techniques in Input-Output Analysis: Table Updating and Structural Analysis[J]. Economic Systems Research, 2004, 16(2): 115-134.

[3] Toh M H. The RAS Approach in Updating Input-output Matrices: An Instrumental Variable Interpretation and Analysis of Structural Change[J]. Economic Systems Research, 1998, 10(1): 63-78.

[4] Liu X J, Liu H X, Fang J F. Higher Quantitative Economics[M]. Beijing: Science Press, 2018: 64-69. (in Chinese)

[5] Paelinck J, Waelbroeck J. Etude Empirique Sur L' évolution Des Coefficients Input-output[J]. JORBEL-Belgian Journal of Operations Research, Statistics, and Computer Science, 1963, 4(1): 3-11.

[6] Gilchrist D A, St Louis L V. Completing Input-output Tables Using Partial Information: with an Application to Canadian Data[J]. Economic Systems Research, 1999, 11(2): 185-193.

[7] Oosterhaven J. GRAS Versus Minimizing Absolute and Squared Differences: a Comment[J]. Economic Systems Research, 2005, 17(3): 327-331.

[8] Chen R H. A Social Accounting Matrix Update Method Combining GRAS and Shorter Squared Distance[J]. Statistics & Decision, 2012(10): 70-72. (in Chinese)

[9] Temursho U, Oosterhaven J, Cardenete M A. A multi-regional Generalized RAS Updating Technique[J]. Spatial Economic Analysis, 2020(2): 1-16.

[10] Lenzen M, Gallego B, Wood R. Matrix Balancing under Conflicting Information[J]. Economic Systems Research, 2009, 21(1): 23-44.

[11] Zhou N N, Li B Y. Research on DRAS Method in Prediction of Fund Flow Matrix Table[J]. Statistics & Information Forum, 2013, 28(07): 15-21. (in Chinese)

[12] Holý V, Šafr K. Disaggregating Input-output Tables by the Multidimensional RAS Method[J]. arXiv preprint arXiv: 1704.07814, 2017.

[13] Lahr M L, Mesnard L D. Biproportional Techniques in Input-Output Analysis: Table Updating and Structural Analysis[J]. Economic Systems Research, 2004, 16: 115-134.

[14] Mesnard L, Miller R E. A Note on Added Information in the RAS Procedure: Reexamination of Some Evidence[J]. Journal of Regional Science, 2006, 46(3): 517-528.

[15] Chen X K, Yang C H. Input-output Analysis[M]. Beijing: Science Press, 2011: 107-113, 117. (in Chinese)

Note: Figure translations are in progress. See original paper for figures.